Badinger, Harald; Egger, Peter

Working Paper

Estimation of higher-order spatial autoregressive panel data error component models

CESifo Working Paper, No. 2556

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Badinger, Harald; Egger, Peter (2009) : Estimation of higher-order spatial autoregressive panel data error component models, CESifo Working Paper, No. 2556, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/26601

Estimation of Higher-Order Spatial Autoregressive Panel Data Error Component Models

HARALD BADINGER
PETER EGGER

CESIFO WORKING PAPER NO. 2556
CATEGORY 10: EMPIRICAL AND THEORETICAL METHODS
FEBRUARY 2009

CESifo *Working Paper No. 2556*

Estimation of Higher-Order Spatial Autoregressive Panel Data Error Component Models

Abstract

This paper develops an estimator for higher-order spatial autoregressive panel data error component models with spatial autoregressive disturbances, SARAR(*R*,*S*). We derive the moment conditions and optimal weighting matrix without distributional assumptions for a generalized moments (GM) estimation procedure of the spatial autoregressive parameters of the disturbance process and define a generalized two-stages least squares estimator for the regression parameters of the model. We prove consistency of the proposed estimators, derive their joint asymptotic distribution, and provide Monte Carlo evidence on their small sample performance.

JEL Code: C13, C21, C23.

Keywords: higher-order spatial dependence, generalized moments estimation, two-stages least squares, asymptotic statistics.

Harald Badinger
Department of Economics
Vienna University of Economics and Business Administration
Althanstrasse 39 - 45
1090 Vienna
Austria
harald.badinger@wu-wien.ac.at

Peter Egger
Ifo Institute for Economic Research and Center for Economic Studies at the University of Munich
Poschingerstrasse 5
81679 Munich
Germany
egger@ifo.de

December 2008

I. Introduction

This paper considers the estimation of panel data models with higher-order spatially autocorrelated error components and spatially autocorrelated dependent variables. Spatial interactions in data may originate from various sources such as strategic interaction between jurisdictions (to attract firms or other mobile agents) and firms (in their price, quantity, or quality setting) or general equilibrium effects which disseminate with spatial decay due to their transmission through trade flows, migration, or input-output relationships.[1] Data sets used in empirical studies often share three features: first, they are available in the form of panel data, with a large cross-sectional and a small time series dimension; second, spatial interactions of various kinds co-exist – such as geography-related, trade-related, migration-related interactions – or the decay function of a single spatial interaction is unknown; third, it is unclear whether spatial interactions are local – and affect only immediate neighbors – or global – and affect second third and other neighbors with repercussions. The estimator proposed here addresses the mentioned three features in a unified framework. It allows for panel data with a fixed but arbitrary number of channels or decay segments of spatial interaction in both the error components and the dependent variable, referred to as SARAR(*R*,*S*).

In developing the estimator, we build on the SARAR(1,1) generalized moments (GM) framework in Kelejian and Prucha (1999) for a single cross-section and the SARAR(0,1) model in Kapoor, Kelejian, and Prucha (2007) for panel data error components models. Obvious advantages of the GM framework is that it does not rely on distributional assumptions and that it can be applied to large data-sets without imposing any restrictions on the matrices of spatial interdependence. We derive GM estimators for the spatial regressive parameters of the disturbance process based on alternative weighting schemes for the moments. We then define a feasible generalized two-stages least squares (FGTSLS) estimator for the model's regression parameters. We determine asymptotic properties of the estimators for the case where the time dimension of the panel and the number of spatial interactions is fixed while the number of cross-sectional units approaches infinity. In particular, we prove that the proposed GM and FGTSLS procedures obtain consistent estimators of the model parameters and we derive their joint asymptotic distribution.

[1] See Cliff and Ord (1973, 1981), Anselin (1988), and Cressie (1993) for classic references about spatial econometric models in general. Recent theoretical contributions of spatial panel data models include Baltagi, Song, and Koh (2003), Baltagi and Li (2004), Baltagi, Song, Jung, and Koh (2007), Kapoor, Kelejian, and Prucha (2007), Korniotis (2008), Baltagi, Egger, and Pfaffermayr (2008), and Lee and Yu (2008). Recent applications of spatial panel data models include Druska and Horrace (2004), Arbia, Basile, and Piras (2005), Egger, Pfaffermayr, and Winner (2005), Baltagi, Egger, and Pfaffermayr (2007), and Badinger and Egger (2008a).

The remainder of the paper is organized as follows. Section II introduces the basic model specification and some notation. Section III proposes GM estimators of the parameters of spatial dependence in the error components based on alternative weighting schemes of the moments. Section IV derives a two-stages least-squares routine to estimate the regression parameters of the model and derives the asymptotic variance-covariance matrix of all model parameters. The latter enables Wald tests on the structure and decay of spatial interactions in the SARAR(*R*,*S*) model. Section V presents the results of a Monte Carlo simulation exercise. Section VI summarizes our main findings and concludes. The appendix sketches the proofs of consistency and the asymptotic distribution of the model parameters, whereas the full details of the proofs are relegated to a technical appendix.

II. Basic Model Specification and Notation

The basic set-up of the error components model with spatially correlated error terms represents a generalization of Kapoor, Kelejian, and Prucha (2007), who consider a panel data error components model with nonstochastic explanatory variables and first-order spatial autoregressive disturbances, i.e., a SAR(1) model. The present paper delivers the following contributions. First, we allow for an *R*-th order spatial autoregressive process in the dependent variable and an *S*-th order spatial process in the disturbances cum error components, i.e., we consider a SARAR(*R*,*S*) panel data error components model. As we show below, this also covers the case of endogenous explanatory variables other than spatial lags of the dependent variable.[2] Second, we prove consistency of proposed generalized moments (GM) estimators of the model parameters and derive their joint asymptotic distribution. In particular, we also relax the normality assumption used, for instance, in Kapoor, Kelejian, and Prucha (2007) to obtain a simplified version of the optimal weighting matrix for the moment conditions. Third, we provide some Monte Carlo evidence with a special emphasis on the spatial model parameter point estimates and the rejection probabilities of Wald tests of the SARAR(*R*,*S*) model against interesting alternatives such as the SARAR(1,1), SARAR(0,*S*), SARAR(*R*,0), and the non-spatial model.

The basic model comprises $i = 1,...,N$ cross-sectional units and $t = 1,...,T$ time periods. For time period *t*, the model reads

$$\mathbf{y}_N(t) = \mathbf{X}_N(t)\boldsymbol{\beta}_N + \sum_{r=1}^{R} \lambda_{r,N} \mathbf{W}_{r,N} \mathbf{y}_N(t) + \mathbf{u}_N(t), \tag{1a}$$

or

[2] See Lee and Liu (2008) and Badinger and Egger (2008b) for cross-sectional spatial models with a SARAR(*R*,*S*) process.

$$\mathbf{y}_N(t) = \mathbf{Z}_N(t)\boldsymbol{\delta}_N + \mathbf{u}_N(t), \tag{1b}$$

where $\mathbf{y}_N(t)$ is an $N \times 1$ vector with cross-sectional observations of the dependent variable in year *t*, $\mathbf{X}_N(t)$ is an $N \times K$ matrix of observations on K non-stochastic explanatory variables, i.e., $\mathbf{X}_N(t) = [\mathbf{x}_{1,N}(t),...,\mathbf{x}_{K,N}(t)]$ with each $N \times 1$ vector $\mathbf{x}_{k,N}(t)$ denoting the observations on the *k*-th explanatory variable. The structure of spatial dependence in $\mathbf{y}_N(t)$ is determined by the time-invariant $N \times N$ matrices $\mathbf{W}_{r,N}$, $r = 1,...,R$, whose elements $w_{ij,r,N}$ are assumed to be known and will often (but need not) be specified as a decreasing function of geographical distance between the cross-sectional units *i* and *j*. The expression $\overline{\mathbf{y}}_{r,N}(t) = \mathbf{W}_{r,N}\mathbf{y}_N(t)$ is referred to as the *r*-th spatial lag of $\mathbf{y}_N$. The specification of a higher-order process allows the strength of spatial interdependence in the dependent variable (reflected in the spatial autoregressive parameters $\lambda_{r,N}$, $r = 1,...,R$) to vary across a fixed number of R subsets of relations between cross-sectional units.

In equation (1b), the $N \times (K + R)$ design matrix is given by $\mathbf{Z}_N(t) = [\mathbf{X}_N(t), \overline{\mathbf{Y}}_N(t)]$, with $\overline{\mathbf{Y}}_N(t) = [\overline{\mathbf{y}}_{1,N}(t),...,\overline{\mathbf{y}}_{R,N}(t)]$, and $\boldsymbol{\delta}_N = (\boldsymbol{\beta}'_N, \boldsymbol{\lambda}'_N)'$, where the $K \times 1$ parameter vector of the exogenous variables is given by $\boldsymbol{\beta}_N = (\beta_{1,N},...,\beta_{K,N})'$ and the $R \times 1$ vector of spatial autoregressive parameters of $\mathbf{y}_N$ is defined as $\boldsymbol{\lambda}_N = (\lambda_{1,N},...,\lambda_{R,N})'$.

The $N \times 1$ vector of error terms $\mathbf{u}_N(t) = [u_{1,N}(t),...,u_{N,N}(t)]'$ is assumed to follow a spatial autoregressive process given by

$$\mathbf{u}_N(t) = \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N}\mathbf{u}_N(t) + \boldsymbol{\varepsilon}_N(t), \tag{1c}$$

$$\boldsymbol{\varepsilon}_N(t) = \boldsymbol{\mu}_N + \mathbf{v}_N(t), \tag{1d}$$

where $\rho_{m,N}$ and $\mathbf{M}_{m,N}$ denote the time-invariant, unknown parameters and the known $N \times N$ matrix of spatial interdependence, respectively. The structure of spatial correlation in the disturbances is determined by the S different, time-invariant $N \times N$ matrices $\mathbf{M}_{m,N}$. As in equation (1a), the specification of a higher-order process allows the strength of spatial interdependence in the disturbances (reflected in the parameters $\rho_{m,N}$, $m = 1,...,S$) to vary across a fixed number of S subsets of relations between cross-sectional units. The expression $\overline{\mathbf{u}}_{m,N}(t) = \mathbf{M}_{m,N}\mathbf{u}_N(t)$ is referred to as the *m*-th spatial lag of $\mathbf{u}_N$. The $S \times 1$ vector of the spatial autoregressive parameters of $\mathbf{u}_N(t)$ is defined as $\boldsymbol{\rho}_N = (\rho_{1,N},...,\rho_{S,N})'$.

Finally, the $N \times 1$ vector of error terms $\boldsymbol{\varepsilon}_N(t)$ consists of two components, $\boldsymbol{\mu}_N$ and $\mathbf{v}_N(t)$. As indicated by the notation, $\boldsymbol{\mu}_N$ is time-invariant while $\mathbf{v}_N(t)$ is not. The typical elements of $\boldsymbol{\varepsilon}_N(t)$ and $\mathbf{v}_N(t)$ are the scalars $\varepsilon_{it,N}$ and $v_{it,N}$, respectively, and the $N \times 1$ vector of unit-specific error components is given by $\boldsymbol{\mu}_N = (\mu_{1,N},...,\mu_{N,N})'$.

Stacking observations for all time periods such that t is the slow index and i is the fast index with all vectors and matrices, the model reads

$$\mathbf{y}_N = \mathbf{X}_N\boldsymbol{\beta}_N + \overline{\mathbf{Y}}_N\boldsymbol{\lambda}_N + \mathbf{u}_N, \tag{2a}$$

or

$$\mathbf{y}_N = \mathbf{Z}_N\boldsymbol{\delta}_N + \mathbf{u}_N, \tag{2b}$$

with the $NT \times K$ regressor matrix $\mathbf{X}_N = [\mathbf{X}'_N(1),...,\mathbf{X}'_N(T)]'$, and $\overline{\mathbf{Y}}_N = (\overline{\mathbf{y}}_{1,N},...,\overline{\mathbf{y}}_{R,N})$, where $\mathbf{y}_{r,N} = [\mathbf{y}'_{r,N}(1),...,\mathbf{y}'_{r,N}(T)]'$ is the $NT \times 1$ vector of observations on the *r*-th spatial lag of the dependent variable $\overline{\mathbf{y}}_{r,N}$. The $NT \times 1$ vector of disturbances $\mathbf{u}_N = [\mathbf{u}'_N(1),...,\mathbf{u}'_N(T)]$ for the spatial autoregressive process of order S is given by

$$\mathbf{u}_N = \sum_{m=1}^{S} \rho_{m,N}(\mathbf{I}_T \otimes \mathbf{M}_{m,N})\mathbf{u}_N + \boldsymbol{\varepsilon}_N, \tag{2c}$$

where $\mathbf{I}_T$ is an identity matrix of dimension $T \times T$. The $NT \times 1$ vector $\boldsymbol{\varepsilon}_N = [\boldsymbol{\varepsilon}'_N(1),...,\boldsymbol{\varepsilon}'_N(T)]'$ is specified as

$$\boldsymbol{\varepsilon}_N = (\mathbf{e}_T \otimes \mathbf{I}_N)\boldsymbol{\mu}_N + \mathbf{v}_N, \tag{3a}$$

where $\mathbf{e}_T$ is a unit vector of dimension $T \times 1$ and $\mathbf{I}_N$ is an identity matrix of dimension $N \times N$. In light of (2c), the error term can also be written as

$$\boldsymbol{\varepsilon}_N = \mathbf{u}_N - \sum_{m=1}^{S} \rho_{m,N}(\mathbf{I}_T \otimes \mathbf{M}_{m,N})\mathbf{u}_N = \mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})\mathbf{u}_N. \tag{3b}$$

It follows that

$$\mathbf{u}_N = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N, \tag{4a}$$

and

$$\mathbf{y}_N = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{r=1}^{R} \lambda_{r,N} \mathbf{W}_{r,N})^{-1}]\mathbf{X}_N(t)\boldsymbol{\beta}_N + [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{r=1}^{R} \lambda_{r,N} \mathbf{W}_{r,N})^{-1}]\mathbf{u}_N, \quad (4b)$$

The following assumptions are maintained throughout this paper.

Assumption 1.
Let T be a fixed positive integer. (a) For all $1 \le t \le T$ and $1 \le i \le N, N \ge 1$, the error components $v_{it,N}$ are identically and (mutually) independently distributed with $E(v_{it,N}) = 0$, $E(v_{it,N}^2) = \sigma_v^2$, where $0 < \sigma_v^2 < b_v < \infty$, and $E|v_{it,N}|^{4+\eta} < \infty$ for some $\eta > 0$. (b) For all $1 \le i \le N, N \ge 1$, the unit-specific error components $\mu_{i,N}$ are identically and (mutually) independently distributed with $E(\mu_{i,N}) = 0$, $E(\mu_{i,N}^2) = \sigma_\mu^2$, where $0 < \sigma_\mu^2 < b_\mu < \infty$, and $E|\mu_{i,N}|^{4+\eta} < \infty$ for some $\eta > 0$. (c) The processes $\{v_{it,N}\}$ and $\{\mu_{i,N}\}$ are independent of each other. Assumption 1 as maintained here is slightly stronger than that in Kapoor, Kelejian, and Prucha (2007), since it requires not only the fourth but also the $(4+\eta)$-th moments of the error components to be finite for some $\eta > 0$. This is required for the central limit theorem of Kelejian and Prucha (2008) to apply, which will be used to derive the asymptotic distribution of the parameter estimates in Section III.

Assumption 1 implies that

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = \sigma_\mu^2 + \sigma_v^2 \text{ for } i = j \text{ and } t = s, \quad (5a)$$

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = \sigma_\mu^2 \text{ for } i = j \text{ and } t \ne s, \quad (5b)$$

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = 0, \text{ otherwise.} \quad (5c)$$

As a consequence, the variance-covariance matrix of the stacked error term $\boldsymbol{\varepsilon}_N$ reads

$$\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N} = E(\boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}_N') = \sigma_\mu^2(\mathbf{J}_T \otimes \mathbf{I}_N) + \sigma_v^2 \mathbf{I}_{NT}, \quad (6a)$$

where $\mathbf{J}_T = \mathbf{e}_T \mathbf{e}_T'$ is a $T \times T$ matrix with unitary elements and $\mathbf{I}_{NT}$ is an identity matrix of dimension $NT \times NT$. Equation (6a) can also be written as

$$\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N} = \sigma_v^2 \mathbf{Q}_{0,N} + \sigma_1^2 \mathbf{Q}_{1,N}, \quad (6b)$$

where $\sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2$. The two matrices $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$, which are central to the estimation of error component models and the moment conditions of the GM estimator, are defined as

$$\mathbf{Q}_{0,N} = (\mathbf{I}_T - \frac{\mathbf{J}_T}{T}) \otimes \mathbf{I}_N , \tag{7}$$

$$\mathbf{Q}_{1,N} = \frac{\mathbf{J}_T}{T} \otimes \mathbf{I}_N . \tag{8}$$

Notice that $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ are both of order $NT \times NT$, symmetric, idempotent, orthogonal to each other, and sum up to $\mathbf{I}_{NT}$.[3]

Assumption 2.

(a) All diagonal elements of the matrices $\mathbf{W}_{r,N}$, $r = 1,...,R$, and $\mathbf{M}_{s,N}$, $s = 1,...,S$, are zero.

(b) The admissible parameter space is restricted as follows:

$\lambda_{r,N} \in (-\underline{a}_N^{\lambda_r}, \overline{a}_N^{\lambda_r})$, with $0 < \underline{a}_N^{\lambda_r}, \overline{a}_N^{\lambda_r} \leq a^{\lambda,r} \leq a^{\lambda} < \infty$, $r = 1, \ldots, R$, and $\sum_{r=1}^{R} |\lambda_{r,N}| < A_\lambda < \infty$.

The first part of Assumption (2b) requires the parameters $\lambda_{r,N}$, $r = 1,...,R$ to be finite. We take a^{λ} such that $a^{\lambda} = \max_{r=1,...,R}(a^{\lambda,r})$ holds; the expression $\mathbf{a}^{\lambda}$ will be used to denote an $R \times 1$ vector with elements a^{λ}. In the second part of Assumption (2b), the scalar A_λ generally depends on the properties of the weights matrices $\mathbf{W}_{r,N}$. For example, with row-normalized matrices $\mathbf{W}_{r,N}$, $r = 1,...,R$, assuming that $A_\lambda = 1$ ensures that $(\mathbf{I}_N - \sum_{r=1}^{R} \lambda_{r,N} \mathbf{W}_{r,N})$ is invertible, as required in Assumption (2c). If the matrices $\mathbf{W}_{r,N}$ are not row-normalized, Assumption (2c) is implied by $A_\lambda = \left(\max_{r=1,...,R} \left\| \mathbf{W}_{r,N} \right\| \right)^{-1}$ for some matrix norm $\| \cdot \|$ (see Horn and Johnson, 1985, p. 301). Analogous assumptions are made for the parameters of the spatial autoregressive error process:

$\rho_{s,N} \in (-\underline{a}_N^{\rho_s}, \overline{a}_N^{\rho_s})$, with $0 < \underline{a}_N^{\rho_s}, \overline{a}_N^{\rho_s} \leq a^{\rho,s} \leq a^{\rho} < \infty$, $s = 1, \ldots, S$, and $\sum_{m=1}^{S} |\rho_{m,N}| < A_\rho < \infty$.

We take a^{ρ} such that $a^{\rho} = \max_{s=1,...,S}(a^{\rho,s})$ holds; the expression $\mathbf{a}^{\rho}$ will be used to denote an $S \times 1$ vector with elements a^{ρ}. As above, with row-normalized matrices $\mathbf{M}_{s,N}$, $s = 1,...,S$, the

[3] Observe that pre-multiplying an $NT \times 1$ vector with $\mathbf{Q}_{0,N}$ transforms its elements into deviations from cross-section specific sample means taken over time, and that pre-multiplying a vector by $\mathbf{Q}_{1,N}$ transforms its elements into cross-section specific sample means. See Remark A.2 in Appendix A for further properties of $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$.

second part of this assumption ensures invertibility of $(\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})$ if $A_\rho = 1$. If the matrices $\mathbf{M}_{s,N}$ are not row-normalized, Assumption (2c) is implied by $A_\rho = \left(\max_{s=1,...,S}\left\|\mathbf{M}_{s,N}\right\|\right)^{-1}$ for some matrix norm $\left\|\cdot\right\|$.

(c) The matrices $(\mathbf{I}_N - \sum_{r=1}^{R}\lambda_{r,N}\mathbf{W}_{r,N})$ and $(\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})$ are nonsingular for $\lambda_r \in (-\underline{a}_N^{\lambda_r}, \overline{a}_N^{\lambda_r})$ and $\rho_s \in (-\underline{a}_N^{\rho_r}, \overline{a}_N^{\rho_r})$, respectively. This ensures that $\mathbf{u}_N$ and $\mathbf{y}_N$ are uniquely identified through equations (4a) and (4b).

Assumption 3.

The row and column sums of the matrices $\mathbf{W}_{r,N}$, $r = 1,...,R$, $\mathbf{M}_{s,N}$, $s = 1,...,S$, $(\mathbf{I}_N - \sum_{r=1}^{R}\lambda_{r,N}\mathbf{W}_{r,N})^{-1}$, and $(\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}$ are bounded uniformly in absolute value. (See Remark A.1 in Appendix A for a definition of row and column sum boundedness.)

In light of Assumptions 1-3 and Remark A.1 in the Appendix, it follows that $E(\mathbf{u}_N) = \mathbf{0}$ and the variance-covariance matrix of $\mathbf{u}_N$ is given by

$$\mathbf{\Omega}_{\mathbf{u},N} = E(\mathbf{u}_N\mathbf{u}'_N) = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\mathbf{\Omega}_{\varepsilon,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})^{-1}], \tag{9a}$$

and

$$E[\mathbf{u}_N(t)\mathbf{u}'_N(t)] = (\sigma_\mu^2 + \sigma_v^2)(\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}(\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})^{-1}. \tag{9b}$$

Note that all variables and parameters except for the variances of the error components are allowed to depend on sample size *N*. Such a specification is consistent, for example, with models where the weights matrix is row-normalized and the number of neighbours of a given cross-sectional unit depends on sample size (see Kapoor, Kelejian, and Prucha, 2007, p. 102) or where the strength of interdependence (in terms of the spatial autoregressive parameters) changes with the number of neighbours. Note that $\mathbf{X}_N$ is allowed to depend on sample size and may thus also contain spatial lags of exogenous variables. As a result, the model specification in equations (1a)-(1c) is fairly general, allowing for higher-order spatial dependence in the dependent variable, the explanatory variables, and the disturbances.

III. GM Estimation of a SAR(*S*) Model

In the following, we consider GM estimators for the spatial autoregressive parameters of the disturbance process in equation (1c) and derive the asymptotic joint distribution of all model parameters.

1. Moment Conditions

With an *S*-th order process (SAR(*S*), with $S>1$), the GM estimators of the parameters $\rho_{1,N},...,\rho_{S,N}$, σ_v^2, and σ_1^2 can be obtained by recognizing that – under Assumptions 1 and 2 – the moment conditions used by Kapoor, Kelejian, and Prucha (2007) hold for each matrix $\mathbf{M}_{s,N}$, $s=1,...,S$. In particular, define for each $\mathbf{M}_{s,N}$, $s=1,...,S$

$$\bar{\boldsymbol{\varepsilon}}_{s,N}=(\mathbf{I}_T\otimes\mathbf{M}_{s,N})\boldsymbol{\varepsilon}_N=(\mathbf{I}_T\otimes\mathbf{M}_{s,N})[\mathbf{u}_N-\sum_{m=1}^{S}\rho_{m,N}(\mathbf{I}_T\otimes\mathbf{M}_{m,N})\mathbf{u}_N]. \tag{10}$$

A word on notation is in order here. In equation (10), subscript *s* has been introduced together with *m* to indicate that, with higher-order spatial processes, $\mathbf{M}_{s,N}$ and $\mathbf{M}_{m,N}$ meet in $\bar{\boldsymbol{\varepsilon}}_{s,N}$. While we will use index s to refer to the matrix $\mathbf{M}_{s,N}$, by which $\boldsymbol{\varepsilon}_N$ is pre-multiplied in equation (10), index m is required for the summation over the terms $\rho_{m,N}\mathbf{M}_{m,N}$.

The moment conditions are then given by

$$\text{M}_\text{a}\quad E[\frac{1}{N(T-1)}\boldsymbol{\varepsilon}_N'\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N]=E[\frac{1}{N(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}\mathbf{v}_N]=\sigma_v^2, \tag{11}$$

$$\text{M}_{1,s}\quad E[\frac{1}{N(T-1)}\bar{\boldsymbol{\varepsilon}}_{s,N}'\mathbf{Q}_{0,N}\bar{\boldsymbol{\varepsilon}}_{s,N}]=E[\frac{1}{N(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{0,N}\mathbf{v}_N]=\sigma_v^2\frac{1}{N}tr(\mathbf{M}_{s,N}'\mathbf{M}_{s,N}),$$

$$\text{M}_{2,s}\quad E[\frac{1}{N(T-1)}\bar{\boldsymbol{\varepsilon}}_{s,N}'\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N]=E[\frac{1}{N(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}')\mathbf{Q}_{0,N}\mathbf{v}_N]=0,$$

$$\text{M}_\text{b}\quad E(\frac{1}{N}\boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N)=E[\frac{1}{N}\boldsymbol{\mu}_N'(\mathbf{e}_T'\mathbf{e}_T\otimes\mathbf{I}_N)\boldsymbol{\mu}_N]+E(\frac{1}{N}\mathbf{v}_N'\mathbf{Q}_{1,N}\mathbf{v}_N)=\sigma_1^2,$$

$$\text{M}_{3,s}\quad E(\frac{1}{N}\bar{\boldsymbol{\varepsilon}}_{s,N}'\mathbf{Q}_{1,N}\bar{\boldsymbol{\varepsilon}}_{s,N})=E[\frac{1}{N}\boldsymbol{\mu}_N'(\mathbf{e}_T'\mathbf{e}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\boldsymbol{\mu}_N]+E[\frac{1}{N}\mathbf{v}_N'\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{1,N}\mathbf{v}_N]$$
$$=\sigma_1^2\frac{1}{N}tr(\mathbf{M}_{s,N}'\mathbf{M}_{s,N}),$$

$$\text{M}_{4,s}\quad E(\frac{1}{N}\bar{\boldsymbol{\varepsilon}}_{s,N}'\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N)=E[\frac{1}{N}\boldsymbol{\mu}_N'(\mathbf{e}_T'\mathbf{e}_T\otimes\mathbf{M}_{s,N}')\boldsymbol{\mu}_N]+E[\frac{1}{N}\mathbf{v}_N'\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}')\mathbf{Q}_{1,N}\mathbf{v}_N]=0,$$

where $\sigma_1^2=\sigma_v^2+T\sigma_\mu^2$. The moment conditions associated with matrices $\mathbf{M}_{s,N}$, $s=1,...,S$, through (10), are indexed with subscripts 1 to 4. The remaining two moment conditions,

which do not depend on s, are denoted as M_a and M_b. For an S-th order process as given by equation (2c), we thus have $(4S+2)$ moment conditions.[4]

Substituting equations (3b), (10), and (1c) into the $4S+2$ moment conditions (11) yields a $(4S+2)$ equation system in $(\rho_{1,N},...,\rho_{S,N},\sigma_v^2,\sigma_1^2)$, which can be written as

$$\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \mathfrak{b}_N = \mathbf{0}, \tag{12}$$

where $\mathfrak{b}_N$ is a $[2S+S(S-1)/2+2]\times 1$ vector, given by

$$\mathfrak{b}_N = (\rho_{1,N},...,\rho_{S,N},\rho_{1,N}^2,...,\rho_{S,N}^2,\rho_{1,N}\rho_{2,N},...,\rho_{1,N}\rho_{S,N},...,\rho_{S-1,N}\rho_{S,N},\sigma_v^2,\sigma_1^2)',$$

i.e., $\mathfrak{b}_N$ contains S linear terms $\rho_{m,N}$, $m=1,...,S$, S quadratic terms $\rho_{m,N}^2$, $m=1,...,S$, $S(S-1)/2$ cross products $\rho_{m,N}\rho_{l,N}$, $m=1,...,S-1, l=m+1,...,S$, as well as σ_v^2 and σ_1^2. For later reference, we define the $(S+2)\times 1$ vector of all parameters as $\boldsymbol{\theta}_N = (\boldsymbol{\rho}_N', \sigma_v^2, \sigma_1^2)' = (\rho_{1,N},...,\rho_{S,N},\sigma_v^2,\sigma_1^2)'$.

$\boldsymbol{\gamma}_N$ is a $(4S+2)\times 1$ vector with elements $[\gamma_{i,N}]$, $i=1,...,(4S+2)$, and $\boldsymbol{\Gamma}_N$ is a $(4S+2)\times[2S+S(S-1)/2+2]$ matrix with elements $[\gamma_{i,j,N}]$, $i=1,...,(4S+2)$, $j=1,...,[2S+S(S-1)/2+2]$. The elements $\gamma_{i,N}$ and $\gamma_{i,j,N}$ will be defined below. The row-index of the elements $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$ will be chosen such that the equation system (12) has the following order. The first four rows correspond to the moment restrictions $M_{1,1}$ to $M_{4,1}$ associated with matrix $\mathbf{M}_{1,N}$ through (10); rows five to eight correspond to $M_{1,2}$ to $M_{4,2}$ associated with matrix $\mathbf{M}_{2,N}$, and so forth; rows $(S-4)$ to $4S$ correspond to the $M_{1,S}$ to $M_{4,S}$ associated with the matrix $\mathbf{M}_{S,N}$. Finally, rows $(4S+1)$ and $(4S+2)$ correspond to the moment conditions M_a and M_b, respectively, which do not depend on s.

[4] Notice that further moment conditions are available through pre- and post-multiplying $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ with $\bar{\boldsymbol{\varepsilon}}_{r,N}'$ and $\bar{\boldsymbol{\varepsilon}}_{s,N}$, $r \neq s$, respectively, in moment conditions M_1 and M_3. The associated efficiency gain will depend on the properties of the weights matrices. If the two involved weights matrices are orthogonal, the corresponding moment condition is trivially satisfied and does not add any information. For the sake of brevity, we use the moment conditions given in (11), and leave an assessment of the potential efficiency gains from exploiting further moment conditions for future research.

The sample analogue to equation system (12) is given by

$$\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \boldsymbol{\beta}_N = \boldsymbol{\vartheta}_N(\boldsymbol{\theta}_N), \tag{13}$$

where the elements of $\tilde{\boldsymbol{\gamma}}_N$ and $\tilde{\boldsymbol{\Gamma}}_N$ are equal to those of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$ with the expectations operator suppressed and the disturbances $\mathbf{u}_N$ replaced by (consistent) estimates $\tilde{\mathbf{u}}_N$.

GM estimates of the parameters $\rho_{1,N,}...,\rho_{S,N}$, σ_v^2 and σ_1^2 are then obtained as the solution to

$$\underset{\rho_1,\rho_2,..,\rho_S,\sigma_v^2,\sigma_1^2}{\arg\min} [(\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \boldsymbol{\beta}_N)' \tilde{\boldsymbol{\Theta}}_N (\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \boldsymbol{\beta}_N)] = [\boldsymbol{\vartheta}_N(\boldsymbol{\theta}_N)' \tilde{\boldsymbol{\Theta}}_N \boldsymbol{\vartheta}_N(\boldsymbol{\theta}_N)], \tag{14}$$

i.e., the parameter estimates can be obtained from a (weighted) non-linear least squares regression of $\tilde{\boldsymbol{\gamma}}_N$ on the columns of $\tilde{\boldsymbol{\Gamma}}_N$; $\boldsymbol{\vartheta}_N(\boldsymbol{\theta}_N)$ can then be viewed as a vector of regression residuals. The optimal choice of the $(4S+2)\times(4S+2)$ weighting matrix $\boldsymbol{\Theta}_N$ and its estimation will be discussed below.

In the following, we define the elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$, grouped by the corresponding moment conditions. Thereby, we use the following notation:

$$\bar{\mathbf{u}}_{s,N} = (\mathbf{I}_T \otimes \mathbf{M}_{s,N})\mathbf{u}_N, \; s = 1,...,S, \text{ and} \tag{15a}$$

$$\bar{\bar{\mathbf{u}}}_{sm,N} = (\mathbf{I}_T \otimes \mathbf{M}_{s,N})(\mathbf{I}_T \otimes \mathbf{M}_{m,N})\mathbf{u}_N = (\mathbf{I}_T \otimes \mathbf{M}_{s,N}\mathbf{M}_{m,N})\mathbf{u}_N, \; s = 1,...,S, \; m = 1,...,S. \tag{15b}$$

At this point, we need to introduce an index $l = 1,...,S$ for proper definition of the elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$.

Moment condition $M_{1,s}$ delivers $s = 1,...,S$ rows of equation system (12), appearing in rows $4(s-1)+1$ with the following elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$:

$$\gamma_{4(s-1)+1,N} = \frac{1}{N(T-1)} \; E(\bar{\mathbf{u}}'_{s,N}\mathbf{Q}_{0,N}\bar{\mathbf{u}}_{s,N}), \tag{16a}$$

$$\gamma_{4(s-1)+1,m,N} = \frac{2}{N(T-1)} E(\bar{\mathbf{u}}'_{s,N}\mathbf{Q}_{0,N}\bar{\bar{\mathbf{u}}}_{sm,N}), \; m = 1,...,S,$$

$$\gamma_{4(s-1)+1,S+m,N} = -\frac{1}{N(T-1)} E(\bar{\bar{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{0,N}\bar{\bar{\mathbf{u}}}_{sm,N}), \; m = 1,...,S,$$

$$\gamma_{4(s-1)+1,S(m+1)-m(m-1)/2+l-m,N} = -\frac{2}{N(T-1)} E(\bar{\bar{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{0,N}\bar{\bar{\mathbf{u}}}_{sl,N}), \; m = 1,...,S-1, \; l = m+1,...,S,$$

$$\gamma_{4(s-1)+1,2S+S(S-1)/2+1,N} = \frac{1}{N} tr(\mathbf{M}'_{s,N}\mathbf{M}_{s,N}),$$

$$\gamma_{4(s-1)+1,2S+S(S-1)/2+2,N} = 0.$$

Moment condition $M_{2,s}$ consists of $s = 1,...,S$ rows of equation system (12), appearing in rows $4(s-1)+2$ with the following elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$:

$$\gamma_{4(s-1)+2,N} = \frac{1}{N(T-1)} E(\overline{\mathbf{u}}'_{s,N}\mathbf{Q}_{0,N}\mathbf{u}_N), \tag{16b}$$

$$\gamma_{4(s-1)+2,m,N} = \frac{1}{N(T-1)} E(\overline{\overline{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{0,N}\mathbf{u}_N + \overline{\mathbf{u}}'_{s,N}\mathbf{Q}_{0,N}\overline{\mathbf{u}}_{m,N}),\ m = 1,...,S,$$

$$\gamma_{4(s-1)+2,S+m,N} = -\frac{1}{N(T-1)} E(\overline{\overline{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{0,N}\overline{\mathbf{u}}_{m,N}),\ m = 1,...,S,$$

$$\gamma_{4(s-1)+2,S(m+1)-m(m-1)/2+l-m,N} = -\frac{1}{N(T-1)} E(\overline{\overline{\mathbf{u}}}'_{sl,N}\mathbf{Q}_{0,N}\overline{\mathbf{u}}_{m,N} + \overline{\overline{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{0,N}\overline{\mathbf{u}}_{l,N}),\ m = 1,...,S-1,$$

$l = m+1,...,S$,

$$\gamma_{4(s-1)+2,2S+S(S-1)/2+1,N} = 0,$$

$$\gamma_{4(s-1)+2,2S+S(S-1)/2+2,N} = 0.$$

Moment condition $M_{3,s}$ corresponds to $s = 1,...,S$ rows of equation system (12), appearing in rows $4(s-1)+3$ with the following elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$:

$$\gamma_{4(s-1)+3,N} = \frac{1}{N} E(\overline{\mathbf{u}}'_{s,N}\mathbf{Q}_{1,N}\overline{\mathbf{u}}_{s,N}), \tag{16c}$$

$$\gamma_{4(s-1)+3,m,N} = \frac{2}{N} E(\overline{\mathbf{u}}'_{s,N}\mathbf{Q}_{1,N}\overline{\overline{\mathbf{u}}}_{sm,N}),\ m = 1,...,S,$$

$$\gamma_{4(s-1)+3,S+m,N} = -\frac{1}{N} E(\overline{\overline{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{1,N}\overline{\overline{\mathbf{u}}}_{sm,N}),\ m = 1,...,S,$$

$$\gamma_{4(s-1)+3,S(m+1)-m(m-1)/2+l-m,N} = -\frac{2}{N} E(\overline{\overline{\mathbf{u}}}'_{sm,N}\mathbf{Q}_{1,N}\overline{\overline{\mathbf{u}}}_{sl,N}),\ m = 1,...,S-1,\ l = m+1,...,S,$$

$$\gamma_{4(s-1)+3,2S+S(S-1)/2+1,N} = 0,$$

$$\gamma_{4(s-1)+3,2S+S(S-1)/2+2,N} = \frac{1}{N} tr(\mathbf{M}'_{s,N}\mathbf{M}_{s,N}).$$

Moment condition $M_{4,s}$ represents $s = 1,...,S$ rows of equation system (12) appearing in rows $4(s-1)+4$ with the following elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$:

$$\gamma_{4(s-1)+4,N} = \frac{1}{N} E(\overline{\mathbf{u}}'_{s,N} \mathbf{Q}_{1,N} \mathbf{u}_N), \tag{16d}$$

$$\gamma_{4(s-1)+4,m,N} = \frac{1}{N} E(\overline{\overline{\mathbf{u}}}'_{sm,N} \mathbf{Q}_{1,N} \mathbf{u}_N + \overline{\mathbf{u}}'_{s,N} \mathbf{Q}_{1,N} \overline{\mathbf{u}}_{m,N}),\ m = 1,...,S,$$

$$\gamma_{4(s-1)+4,S+m,N} = -\frac{1}{N} E(\overline{\overline{\mathbf{u}}}'_{sm,N} \mathbf{Q}_{1,N} \overline{\mathbf{u}}_{m,N}),\ m = 1,...,S,$$

$$\gamma_{4(s-1)+4,S(m+1)-m(m-1)/2+l-m,N} = -\frac{1}{N} E(\overline{\overline{\mathbf{u}}}'_{sl,N} \mathbf{Q}_{1,N} \overline{\mathbf{u}}_{m,N} + \overline{\overline{\mathbf{u}}}'_{sm,N} \mathbf{Q}_{1,N} \overline{\mathbf{u}}_{l,N}),\ m = 1,...,S-1,\ l = m+1,...,S,$$

$$\gamma_{4(s-1)+4,2S+S(S-1)/2+1,N} = 0,$$

$$\gamma_{4(s-1)+4,2S+S(S-1)/2+2,N} = 0.$$

Moment condition M_a reflects one equation of the system in (12), appearing in row $(4S+1)$ with the following elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$:

$$\gamma_{4S+1,N} = \frac{1}{N(T-1)} E(\mathbf{u}'_N \mathbf{Q}_{0,N} \mathbf{u}_N), \tag{16e}$$

$$\gamma_{4S+1,m,N} = \frac{2}{N(T-1)} E(\overline{\mathbf{u}}'_{m,N} \mathbf{Q}_{0,N} \mathbf{u}_N),\ m = 1,...,S,$$

$$\gamma_{4S+1,S+m,N} = -\frac{1}{N(T-1)} E(\overline{\mathbf{u}}'_{m,N} \mathbf{Q}_{0,N} \overline{\mathbf{u}}_{m,N}),\ m = 1,...,S,$$

$$\gamma_{4S+1,S(m+1)-m(m-1)/2+l-m,N} = -\frac{2}{N(T-1)} E(\overline{\mathbf{u}}'_{m,N} \mathbf{Q}_{0,N} \overline{\mathbf{u}}_{l,N}),\ m = 1,...,S-1,\ l = m+1,...,S,$$

$$\gamma_{4S+1,2S+S(S-1)/2+1,N} = 1,$$

$$\gamma_{4S+1,2S+S(S-1)/2+2,N} = 0.$$

Moment condition M_b is associated with one equation of the system in (12), appearing in row $(4S+2)$ with the following elements of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$:

$$\gamma_{4S+2,N} = \frac{1}{N} E(\mathbf{u}'_N \mathbf{Q}_{1,N} \mathbf{u}_N), \tag{16f}$$

$$\gamma_{4S+2,m,N} = \frac{2}{N} E(\overline{\mathbf{u}}'_{m,N} \mathbf{Q}_{1,N} \mathbf{u}_N),\ m = 1,...,S,$$

$$\gamma_{4S+2,S+m,N} = -\frac{1}{N} E(\overline{\mathbf{u}}'_{m,N} \mathbf{Q}_{1,N} \overline{\mathbf{u}}_{m,N}),\ m = 1,...,S,$$

$$\gamma_{4S+2,S(m+1)-m(m-1)/2+l-m,N} = -\frac{2}{N} E(\overline{\mathbf{u}}'_{m,N} \mathbf{Q}_{1,N} \overline{\mathbf{u}}_{l,N}),\ m = 1,...,S-1,\ l = m+1,...,S,$$

$$\gamma_{4S+2,2S+S(S-1)/2+1,N} = 0,$$

$$\gamma_{4S+2,2S+S(S-1)/2+2,N} = 1.$$

This completes the specification of the elements of the matrices $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$. The similarity of the structure between the expressions resulting from the moment conditions M_a, $M_{1,s}$, and $M_{2,s}$ on the one hand and M_b, $M_{3,s}$, and $M_{4,s}$ on the other hand is apparent. First, they differ by the normalization factor and the corresponding matrix of quadratic forms, $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$, respectively. Second, the rows in (12) associated with M_a, $M_{1,s}$ and $M_{2,s}, s = 1,...,S$, do not depend on σ_1^2 whereas the rows associated with M_b, $M_{3,s}$, and $M_{4,s}$, $s = 1,...,S$, do not depend on σ_v^2. This fact will be used to define an initial GM estimator, which is based on a subset of moment conditions (M_a, $M_{1,s}$, and $M_{2,s}$) only, in order to obtain an estimate of the optimal weighting matrix $\boldsymbol{\Theta}_N$.

For future reference, we define the $(2S+1)\times 1$ vector $\boldsymbol{\gamma}_N^0$ as the sub-vector containing rows s and $(s+1)$, $s = 1,...,S$ and row $(4S+1)$ of $\boldsymbol{\gamma}_N$, corresponding to $M_{1,s}$, $M_{2,s}$, and M_a. Moreover, we define the $(2S+1)\times[2S+S(S-1)/2+1]$ matrix $\boldsymbol{\Gamma}_N^0$ as the sub-matrix containing rows s and $(s+1)$, $s = 1,...,S$, and row $(4S+1)$ of $\boldsymbol{\Gamma}_N$, corresponding to $M_{1,s}$, $M_{2,s}$, and M_a.

Analogously, we define the $(2S+1)\times 1$ vector $\boldsymbol{\gamma}_N^1$ as the sub-vector containing rows $2s$, $(2s+1)$, $s = 1,...,S$, and row $(4S+2)$ of $\boldsymbol{\gamma}_N$, corresponding to $M_{3,s}$, $M_{4,s}$, and M_b. Finally, we define the $(2S+1)\times[2S+S(S-1)/2+1]$ matrix $\boldsymbol{\Gamma}_N^1$ as the sub-matrix containing rows $2s$, $(2s+1)$, $s = 1,...,S$, and $(4S+2)$ of $\boldsymbol{\Gamma}_N$, corresponding to $M_{3,s}$, $M_{4,s}$, and M_b.

2. Definition of GM Estimators

We next define three alternative GM estimators for the spatial autoregressive parameters of the disturbance process given by (1c) and the variances of the error components.[5]

2.1. Initial GM Estimation

The initial GM estimator is a special case of (14), using the identity matrix as weighting matrix $\boldsymbol{\Theta}_N$ and a subset of moment conditions (M_a, $M_{1,s}$ and $M_{2,s}$) only. It is thus based on the the vector $\boldsymbol{\gamma}_N^0$ and the matrix $\boldsymbol{\Gamma}_N^0$. Define $\boldsymbol{\theta}_N^0$ as the corresponding parameter vector that excludes σ_1^2, i.e., $\boldsymbol{\theta}^0 = (\boldsymbol{\rho}_N', \sigma_v^2) = (\rho_{1,N},...,\rho_{S,N},\sigma_v^2)$, and accordingly

$\boldsymbol{\vartheta}_N^0 = (\rho_{1,N},...,\rho_{S,N}, \rho_{1,N}^2,...,\rho_{S,N}^2, \rho_{1,N}\rho_{2,N},...,\rho_{1,N}\rho_{S,N},...,\rho_{S-1,N}\rho_{S,N},\sigma_v^2)'$.

The initial GM estimator is then obtained as the solution to

[5] See Kapoor, Kelejian, and Prucha (2007) for analogous conditions under SARAR(0,1) estimation, assuming only nonstochastic regressors in equation (1a).

$$(\widehat{\rho}_{1,N},...,\widehat{\rho}_{S,N},\widehat{\sigma}^2_{v,N}) = \arg\min\{\vartheta^0_N(\underline{\boldsymbol{\theta}}^0_N)'\vartheta^0_N(\underline{\boldsymbol{\theta}}^0_N), -\mathbf{a}^\rho \le \underline{\boldsymbol{\rho}} \le \mathbf{a}^\rho, \underline{\sigma}^2_{v,N} \in [0,b_v]\}, \tag{17a}$$

$$\text{with } \vartheta^0_N(\underline{\boldsymbol{\theta}}^0) = \vartheta^0_N(\underline{\boldsymbol{\rho}},\underline{\sigma}^2_v) = (\widetilde{\boldsymbol{\gamma}}^0_N - \widetilde{\boldsymbol{\Gamma}}^0_N \underline{\mathbf{b}}^0).$$

Using these initial estimates of $(\rho_{1,N},...,\rho_{S,N})$ and σ^2_v, the parameter σ^2_1 can be estimated from moment condition M_b:

$$\widehat{\sigma}^2_{1,N} = \frac{1}{N}(\widetilde{\mathbf{u}}_N - \sum_{m=1}^{S}\widehat{\rho}_{m,N}\widetilde{\overline{\mathbf{u}}}_{m,N})'\mathbf{Q}_{1,N}(\widetilde{\mathbf{u}}_N - \sum_{m=1}^{S}\widehat{\rho}_{m,N}\widetilde{\overline{\mathbf{u}}}_{m,N}) \tag{17b}$$

$$= \widetilde{\gamma}_{4S+2} - \widetilde{\gamma}_{4S+2,1}\widehat{\rho}_{1,N} - ... - \widetilde{\gamma}_{4S+2,S}\widehat{\rho}_{S,N} - \widetilde{\gamma}_{4S+2,S+1}\widehat{\rho}^2_{1,N}...$$

$$- \widetilde{\gamma}_{4S+2,2S}\widehat{\rho}^2_{S,N} - \widetilde{\gamma}_{4S+2,2S+1}\widehat{\rho}_{1,N}\widehat{\rho}_{2,N} - ... - \widetilde{\gamma}_{4S+2,2S+S(S-1)/2}\widehat{\rho}_{S-1,N}\widehat{\rho}_{S,N}.$$

2.2. Weighted GM Estimation

While the initial GM estimator as defined in (17) is consistent (as will be shown below), it is inefficient. First, it ignores the information contained in moment conditions (M_b, $M_{3,s}$ and $M_{4,s}$).[6] Second, it is well known from the literature on generalized method of moments estimation, that it is optimal to use as weighting matrix the inverse of the (properly normalized) variance-covariance matrix of the moments, evaluated at true parameter values. Denote the optimal weighting matrix, which will be derived in Subsection 3.2, by $\boldsymbol{\Psi}^{-1}_N$. In general, the optimal weighting matrix is unknown and has to be estimated, e.g., using the results from the initial GM estimation. In Subsection 3.3 we derive a consistent estimator of $\boldsymbol{\Psi}^{-1}_N$, referred to as $\widetilde{\boldsymbol{\Psi}}^{-1}_N$. The optimally weighted GM estimator is based on all $(4S+2)$ moment conditions and uses $\widetilde{\boldsymbol{\Theta}}_N = \widetilde{\boldsymbol{\Psi}}^{-1}_N$ as the weighting matrix for the moment conditions. It is defined as:

$$(\widetilde{\rho}_{1,N},...,\widetilde{\rho}_{S,N},\widetilde{\sigma}^2_{v,N},\widetilde{\sigma}^2_{1,N}) = \arg\min\{\vartheta_N(\underline{\boldsymbol{\theta}})'\widetilde{\boldsymbol{\Theta}}_N\vartheta_N(\underline{\boldsymbol{\theta}}), -\mathbf{a}^\rho \le \underline{\boldsymbol{\rho}} \le \mathbf{a}^\rho, \underline{\sigma}^2_v \in [0,b_v], \underline{\sigma}^2_1 \in [0,c]\},$$

$$\text{with } c \ge b_v + Tb_\mu, \text{ and } \vartheta_N(\underline{\boldsymbol{\theta}}) = \vartheta_N(\underline{\boldsymbol{\rho}},\underline{\sigma}^2_v,\underline{\sigma}^2_1) = (\widetilde{\boldsymbol{\gamma}}_N - \widetilde{\boldsymbol{\Gamma}}_N\underline{\mathbf{b}}). \tag{18}$$

As already mentioned, the optimal weighting matrix is derived without distributional assumptions. As a consequence it involves third and fourth moments of the error components

[6] This does not mean that any GM estimator using all moment conditions is necessarily superior. In fact, Kapoor, Kelejian, and Prucha (2007) show in a Monte Carlo study on their SAR(1) model that the initial GM estimator performs much better (in terms of bias and RMSE) than the unweighted GM estimator using all moment conditions. Their results suggest that a proper weighting of the moment conditions, in particular the weighting of moment conditions M_a, M_1, and M_2 relative to moment conditions M_b, M_3, and M_4 is of crucial importance.

$v_{it,N}$ and $\mu_{i,N}$. Kapoor, Kelejian, and Prucha (2008) use the assumption that $\varepsilon_{it,N}$ is normally distributed to obtain a simplified weighting matrix as an approximation of the true optimal weighting matrix. For comparison, we also consider this simplified optimal weighting matrix, which is shown to be a special case of $\mathbf{\Psi}_N^{-1}$ in the Appendix and referred to as $(\mathbf{\Psi}_N^{\circ})^{-1}$. The simplified weighted GM estimator is defined as the weighted GM estimator given in (18), using $\widetilde{\mathbf{\Theta}}_N = (\widetilde{\mathbf{\Psi}}_N^{\circ})^{-1}$.

3. Asymptotic Properties of the GM Estimator for $\boldsymbol{\theta}_N$

3.1 Consistency

In order to prove consistency, the following additional assumptions are required:

Assumption 4.

Assume that $\widetilde{\mathbf{u}}_N - \mathbf{u}_N = \mathbf{D}_N \mathbf{\Delta}_N$, i.e., $\widetilde{u}_{i,N} - u_{i,N} = \mathbf{d}_{i.,N} \mathbf{\Delta}_N$, for $i = 1,...,NT$,[7] where $\mathbf{D}_N$ is an $NT \times P$ matrix, the $1 \times P$ vector $\mathbf{d}_{i.,N}$ denotes the *i*-th row of $\mathbf{D}_N$ and $\mathbf{\Delta}_N$ is a $P \times 1$ vector. Let $d_{ij,N}$ be the *j*-th element of $\mathbf{d}_{i.,N}$. For some $\delta > 0$, we assume that $E\left|d_{ij,N}(t)\right|^{2+\delta} \leq c_d < \infty$, where c_d does not depend on *N*, and that $N^{1/2}\left\|\mathbf{\Delta}_N\right\| = O_p(1)$.

Assumption 4 will be fulfilled in many settings, e.g., if model (1a) contains endogenous variables (such as spatial lags of $\mathbf{y}_N$) and is estimated using two-stages least squares. In that case, $\mathbf{\Delta}_N$ denotes the difference between the parameter estimates and the true parameter values and $\mathbf{d}_{i.,N}$ is the (negative of the) *i*-th row of the design matrix $\mathbf{Z}_N$ (compare Lemma 1 in Subsection 2 of Section IV). Under certain conditions, Assumption 4 will also be satisfied if model (1a) involves a non-linear specification (see Kelejian and Prucha, 2008, p. 12). Finally, observe that Assumption 4 implies that $(NT)^{-1}\sum_{i=1}^{NT}\left\|\mathbf{d}_{i.,N}\right\|^{2+\delta}$ is $O_p(1)$.

Assumption 5.

(a) The smallest eigenvalues of $\mathbf{\Gamma}_N^{0\prime}\mathbf{\Gamma}_N^0$ and $\mathbf{\Gamma}_N^{1\prime}\mathbf{\Gamma}_N^1$ are bounded away from zero, i.e., $\lambda_{\min}(\mathbf{\Gamma}_N^{i\prime}\mathbf{\Gamma}_N^i) \geq \lambda_* > 0$ for i = 1, 2. (b) $\widetilde{\mathbf{\Theta}}_N - \mathbf{\Theta}_N = o_p(1)$, where $\mathbf{\Theta}_N$ are $(4S+2) \times (4S+2)$ nonstochastic, symmetric, positive definite matrices. (c) The largest eigenvalues of $\mathbf{\Theta}_N$ are bounded uniformly from above, and the smallest eigenvalues of $\mathbf{\Theta}_N$ are bounded uniformly away from zero.

---

[7] Note that we use single indexation $i = 1,...,NT$ to refer to the elements of the vectors that are stacked over time periods. (See the remark on notation in Appendix A.)

Assumption 5 implies that the smallest eigenvalues of $\mathbf{\Gamma}_N'\mathbf{\Gamma}_N$ and $\mathbf{\Gamma}_N'\mathbf{\Theta}_N\mathbf{\Gamma}_N$ are bounded uniformly away from zero, ensuring that the true parameter vector $\boldsymbol{\theta}_N$ is identifiable unique. Moreover, by the equivalence of matrix norms, it follows from Assumption 5 that $\mathbf{\Theta}_N$ and $\mathbf{\Theta}_N^{-1}$ are $O(1)$.

Assumptions 1-5 ensure consistency of the GM estimators for $\boldsymbol{\theta}_N = (\boldsymbol{\rho}_N', \sigma_v^2, \sigma_1^2)$. We summarize these results in the following theorems, which are proved in Appendix B.

Theorem 1a. Consistency of Initial GM Estimator $\tilde{\boldsymbol{\theta}}_N^0$

Suppose Assumptions 1-5 hold. Then, provided the optimization space contains the parameter space, the initial GM estimators $\hat{\boldsymbol{\theta}}_N^0 = (\hat{\rho}_{1,N},...,\hat{\rho}_S,\hat{\sigma}_{v,N}^2)'$ defined by (17a), and $\hat{\sigma}_{1,N}^2$, defined by (17b) are consistent for $\rho_{1,N},...,\rho_{S,N}$, σ_v^2, and σ_1^2, i.e.,

$$\hat{\rho}_{s,N} - \rho_{s,N} \xrightarrow{p} 0,\ s = 1,...,S,\ \ \hat{\sigma}_{v,N}^2 - \sigma_v^2 \xrightarrow{p} 0 \text{ , and } \hat{\sigma}_{1,N}^2 - \sigma_1^2 \xrightarrow{p} 0 \text{ as } N \to \infty .$$

Theorem 1b. Consistency of Weighted GM Estimator $\tilde{\boldsymbol{\theta}}_N$

Suppose Assumptions 1-5 hold. Then, provided the optimization space contains the parameter space, the weighted GM estimators $\tilde{\boldsymbol{\theta}}_N(\tilde{\mathbf{\Theta}}_N) = [\tilde{\rho}_{1,N}(\tilde{\mathbf{\Theta}}_N),...,\tilde{\rho}_{S,N}(\tilde{\mathbf{\Theta}}_N),\tilde{\sigma}_{v,N}^2(\tilde{\mathbf{\Theta}}_N),\tilde{\sigma}_{1,N}^2(\tilde{\mathbf{\Theta}}_N)]'$ defined by (18) are consistent for $\rho_{1,N},...,\rho_{S,N},\sigma_v^2$, and σ_1^2, i.e.,

$$\tilde{\rho}_{s,N}(\tilde{\mathbf{\Theta}}_N) - \rho_{s,N} \xrightarrow{p} 0,\ s = 1,...,S,\ \tilde{\sigma}_{v,N}^2(\tilde{\mathbf{\Theta}}_N) - \sigma_v^2 \xrightarrow{p} 0 \text{ , and } \tilde{\sigma}_{1,N}^2(\tilde{\mathbf{\Theta}}_N) - \sigma_1^2 \xrightarrow{p} 0 \text{ as } N \to \infty .$$

This result holds for an arbitrary weighting matrix (that satisfies Assumption 5). Hence, it applies to both the optimally weighted GM estimator defined by (18) with $\tilde{\mathbf{\Theta}}_N = (\tilde{\mathbf{\Psi}}_N)^{-1}$ and the simplified optimally weighted GM estimator $\tilde{\boldsymbol{\theta}}_N^\circ$ with $\tilde{\mathbf{\Theta}}_N = (\tilde{\mathbf{\Psi}}_N^\circ)^{-1}$.

3.2 Asymptotic Distribution of GM Estimator for $\boldsymbol{\theta}_N$

In the following we consider the asymptotic distribution of the optimally weighted GM estimator $\tilde{\boldsymbol{\theta}}_N$. To establish asymptotic normality of $\tilde{\boldsymbol{\theta}}_N = (\tilde{\boldsymbol{\rho}}_N, \tilde{\sigma}_{v,N}^2, \tilde{\sigma}_{1,N}^2)$, we need some additional assumptions.

Assumption 6.

Let $\mathbf{D}_N$ be defined as in Assumption 4, such that $\tilde{\mathbf{u}}_N - \mathbf{u}_N = \mathbf{D}_N\mathbf{\Delta}_N$. For any real $N \times N$ matrix $\mathbf{A}_N$, whose row and column sums are bounded uniformly in absolute value, it holds that $N^{-1}\mathbf{D}_N'\mathbf{A}_N\mathbf{u}_N - N^{-1}E(\mathbf{D}_N'\mathbf{A}_N\mathbf{u}_N) = o_p(1)$.

A sufficient condition for Assumption 6 is, e.g., that the columns of $\mathbf{D}_N$ are of the form $\boldsymbol{\pi}_N + \boldsymbol{\Pi}_N \boldsymbol{\varepsilon}_N$, where the elements of $\boldsymbol{\pi}_N$ are bounded uniformly in absolute value and the row and column sums of $\boldsymbol{\Pi}_N$ are bounded uniformly in absolute value (see Kelejian and Prucha, 2008, Lemma C.2). This will be the case in many applications, e.g., for the model in equation (1a), if $\mathbf{D}_N$ equals (the negative of) matrix $\mathbf{Z}_N$ (compare Lemma 1 in Section IV).

Assumption 7.
Let $\boldsymbol{\Delta}_N$ be defined as in Assumption 4. Then,

$$(NT)^{1/2}\boldsymbol{\Delta}_N = (NT)^{-1/2}\mathbf{T}_N'\boldsymbol{\xi}_N + o_p(1), \text{ with } \mathbf{T}_N = (\mathbf{T}_{v,N}', \mathbf{T}_{\mu,N}')', \ \boldsymbol{\xi}_N = (\mathbf{v}_N', \boldsymbol{\mu}_N')', \text{ i.e.,}$$
$$(NT)^{1/2}\boldsymbol{\Delta}_N = (NT)^{-1/2}\mathbf{T}_{v,N}'\mathbf{v}_N + (NT)^{-1/2}\mathbf{T}_{\mu,N}'\boldsymbol{\mu}_N + o_p(1),$$

where $\mathbf{T}_N$ is an $(NT+N)\times P$-dimensional real nonstochastic matrix whose elements are bounded uniformly in absolute value; $\mathbf{T}_{v,N}$ is of dimension $(NT \times P)$ and $\mathbf{T}_{\mu,N}$ is of dimension $(N \times P)$. As remarked above, $\boldsymbol{\Delta}_N$ typically denotes the difference between the parameter estimates and the true parameter values. Assumption 7 will be satisfied by many estimators. In Section IV, we verify that it holds if the model in equation (1a) is estimated by two-stages least squares (TSLS) or feasible generalized two-stages least squares (FGTSLS).

The limiting distribution of the GM estimator of $\boldsymbol{\theta}_N$ will be seen to depend on (the inverse of) the matrix $\mathbf{J}_N'\boldsymbol{\Theta}_N\mathbf{J}_N$ and the variance-covariance matrix of a vector of quadratic forms in $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$, denoted as $\mathbf{q}_N$. We consider each of these expressions in the following. The $(4S+2)\times(S+2)$ matrix $\mathbf{J}_N$ of derivatives of the $(4S+2)\times 1$ vector of moment conditions in (11) is given by

$$\mathbf{J}_N(\boldsymbol{\theta}_N) = \frac{\partial(\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N\mathbf{b}_N)}{\partial\boldsymbol{\theta}'} = (j_{i,1,N},...,j_{i,S,N},j_{i,\sigma_v,N},j_{i,\sigma_1,N}), \text{ with} \tag{19a}$$
$$j_{i,s,N} = \frac{\partial(\boldsymbol{\gamma}_{i.,N} - \boldsymbol{\Gamma}_{i.,N}\mathbf{b}_N)}{\partial\rho_s}, \ i=1,...,(4S+2), \ s=1,...,S,$$
$$j_{i,\sigma_v,N} = \frac{\partial(\boldsymbol{\gamma}_{i.,N} - \boldsymbol{\Gamma}_{i.,N}\mathbf{b}_N)}{\partial\sigma_v}, \ i=1,...,(4S+2),$$
$$j_{i,\sigma_1,N} = \frac{\partial(\boldsymbol{\gamma}_{i.,N} - \boldsymbol{\Gamma}_{i.,N}\mathbf{b}_N)}{\partial\sigma_1}, \ i=1,...,(4S+2),$$

where $\boldsymbol{\gamma}_{i.,N}$ and $\boldsymbol{\Gamma}_{i.,N}$ denote the *i*-th row of $\boldsymbol{\gamma}_N$ and $\boldsymbol{\Gamma}_N$ respectively.

Using $\frac{\partial \boldsymbol{\gamma}_N}{\partial \boldsymbol{\theta}'} = \mathbf{0}$ and ignoring the negative sign, we have

$$\mathbf{J}_N(\boldsymbol{\rho}) = \frac{\partial}{\partial \boldsymbol{\theta}'} \boldsymbol{\Gamma}_N \mathfrak{b}_N = \boldsymbol{\Gamma}_N \mathfrak{B}_N, \tag{19b}$$

where $\boldsymbol{\Gamma}_N$ is defined above and of dimension $(4S+2)\times[2S+S(S-1)/2+2]$ and $\mathfrak{B}_N$ is a $[2S+S(S-1)/2+2]\times(S+2)$ matrix of the form

$$\mathfrak{B}_N = (\mathfrak{B}_1, \mathfrak{B}'_{2,N}, \mathfrak{B}'_{3,N}, \mathfrak{B}'_{4,N})', \tag{20a}$$

with $\mathfrak{B}_1 = (\mathbf{I}_S, \mathbf{0}_{S\times 2})$, $\mathfrak{B}_{2,N} = [diag_{s=1}^{S}(2\rho_{s,N}), \mathbf{0}_{S\times 2})]$, and $\mathfrak{B}_{3,N} = (\mathfrak{B}'_{3,1,N}, ..., \mathfrak{B}'_{3,S-1,N})'$ is an $S(S-1)/2 \times S$ matrix, consisting of $(S-1)$ vertically arranged blocks $\mathfrak{B}_{3,m,N}$, $m = 1, ..., (S-1)$, which have the following structure:

$$\mathfrak{B}_{3,m,N} = (\mathfrak{C}_{m,N}, \mathfrak{d}_{m,N}, \mathfrak{E}_{m,N}, \mathbf{0}_{(S-m)\times 2}), \tag{20b}$$

where $\mathfrak{C}_{m,N}$ is a $(S-m)\times(m-1)$ matrix of zeros,[8] $\mathfrak{d}_{m,N}$ is a $(S-m)\times 1$ vector, defined as $\mathfrak{d}_{m,N} = (\rho_{m+1,N}, ..., \rho_{S,N})$, and $\mathfrak{E}_{m,N} = \rho_{m,N}\mathbf{I}_{S-m}$. Finally, $\mathfrak{B}_{4,N}$ is a $2\times(S+2)$ matrix, defined as

$$\mathfrak{B}_{4,N} = \begin{bmatrix} \mathbf{0}_{1\times S}, 1, 0 \\ \mathbf{0}_{1\times S+1}, 1 \end{bmatrix}. \tag{20c}$$

For later reference, note that $\mathfrak{B}_N$ has full column rank $(S+2)$; as a consequence, the $(S+2)\times(S+2)$ matrix $\mathfrak{B}'_N\mathfrak{B}_N$ is positive definite (see, e.g., Greene, 2003, p. 835).

We next consider the vector $\mathbf{q}_N$ and its limiting distribution. First, define $\mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N)$ as the $(4S+2)\times 1$ vector of sample moments as given by (11) with the expectation operator suppressed, evaluated at the true parameter values, and ignoring the deterministic constants:

[8] I.e., there is no block $\mathfrak{C}_{1,N}$ in $\mathfrak{B}_{3,1,N}$.

$$\mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N) = N^{-1}\begin{bmatrix} \tilde{\mathbf{u}}_N'\mathbf{C}_{1,1,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{2,1,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{3,1,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{4,1,N}\tilde{\mathbf{u}}_N \\ \vdots \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{1,S,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{2,S,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{3,S,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{4,S,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{a,N}\tilde{\mathbf{u}}_N \\ \tilde{\mathbf{u}}_N'\mathbf{C}_{b,N}\tilde{\mathbf{u}}_N \end{bmatrix}, \tag{21}$$

where

$$\mathbf{C}_{1,s,N} = \frac{1}{(T-1)}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})]\mathbf{Q}_{0,N}(\mathbf{I}_T \otimes \mathbf{M}'_{s,N}\mathbf{M}_{s,N})\mathbf{Q}_{0,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})],$$

$$\mathbf{C}_{2,s,N} = \frac{1}{2(T-1)}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})]\mathbf{Q}_{0,N}[\mathbf{I}_T \otimes (\mathbf{M}_{s,N} + \mathbf{M}'_{s,N})]\mathbf{Q}_{0,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})],$$

$$\mathbf{C}_{3,s,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})]\mathbf{Q}_{1,N}(\mathbf{I}_T \otimes \mathbf{M}'_{s,N}\mathbf{M}_{s,N})\mathbf{Q}_{1,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})],$$

$$\mathbf{C}_{4,s,N} = \frac{1}{2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})]\mathbf{Q}_{1,N}[\mathbf{I}_T \otimes (\mathbf{M}_{s,N} + \mathbf{M}'_{s,N})]\mathbf{Q}_{1,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})],$$

$$\mathbf{C}_{a,N} = \frac{1}{(T-1)}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})]\mathbf{Q}_{0,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})],$$

$$\mathbf{C}_{b,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})]\mathbf{Q}_{1,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]. \tag{22}$$

By Assumption 3 and Remark A.1 in Appendix A, the row and column sums of the symmetric $NT \times NT$ matrices $\mathbf{C}_{p,s,N}$, $p = 1,...,4$, $s = 1,...,S$, $\mathbf{C}_{a,N}$, and $\mathbf{C}_{b,N}$ are bounded uniformly in absolute value. Also, note that $\mathbf{C}_{3,s,N}$, $\mathbf{C}_{4,s,N}$, $\mathbf{C}_{b,N}$ differ from $\mathbf{C}_{1,s,N}$, $\mathbf{C}_{2,s,N}$, $\mathbf{C}_{a,N}$ only by the normalization and the use of $\mathbf{Q}_{1,N}$ versus $\mathbf{Q}_{0,N}$.

In light of (21) and Lemma B.1 (see Appendix B), the elements of $N^{1/2}\mathbf{q}_N(\boldsymbol{\rho}_N,\boldsymbol{\Delta}_N)$ can be expressed as

$$N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N)=\begin{bmatrix} N^{-1/2}\mathbf{u}_N'\mathbf{C}_{1,1,N}\mathbf{u}_N+\boldsymbol{\alpha}_{1,1,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{2,1,N}\mathbf{u}_N+\boldsymbol{\alpha}_{2,1,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{3,1,N}\mathbf{u}_N+\boldsymbol{\alpha}_{3,1,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{4,1,N}\mathbf{u}_N+\boldsymbol{\alpha}_{4,1,N}'N^{1/2}\boldsymbol{\Delta}_N \\ \cdot \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{1,S,N}\mathbf{u}_N+\boldsymbol{\alpha}_{1,S,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{2,S,N}\mathbf{u}_N+\boldsymbol{\alpha}_{2,S,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{3,S,N}\mathbf{u}_N+\boldsymbol{\alpha}_{3,S,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{4,S,N}\mathbf{u}_N+\boldsymbol{\alpha}_{4,S,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{a,N}\mathbf{u}_N+\boldsymbol{\alpha}_{a,N}'N^{1/2}\boldsymbol{\Delta}_N \\ N^{-1/2}\mathbf{u}_N'\mathbf{C}_{a,N}\mathbf{u}_N+\boldsymbol{\alpha}_{b,N}'N^{1/2}\boldsymbol{\Delta}_N \end{bmatrix}+o_p(1), \tag{23}$$

where $\boldsymbol{\alpha}_{p,s,N}=2N^{-1}E(\mathbf{D}_N'\mathbf{C}_{p,s,N}\mathbf{u}_N)$, $p=1,...,4$, $s=1,...,S$, $\boldsymbol{\alpha}_{a,N}=2N^{-1}E(\mathbf{D}_N'\mathbf{C}_{a,N}\mathbf{u}_N)$, and $\boldsymbol{\alpha}_{b,N}=2N^{-1}E(\mathbf{D}_N'\mathbf{C}_{b,N}\mathbf{u}_N)$. By Lemma B.1 the elements of the $P\times 1$ vectors $\boldsymbol{\alpha}_{p,s,N}$, $p=1,...,4$, $s=1,...,S$, $\boldsymbol{\alpha}_{a,N}$ and $\boldsymbol{\alpha}_{b,N}$ are bounded uniformly in absolute value.

Using (22), (3c), Assumption 7, and $\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N=\mathbf{Q}_{0,N}\mathbf{v}_N$ we obtain:

$$N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N)=N^{-1/2}\begin{bmatrix} \frac{1}{(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{0,N}\mathbf{v}_N+\mathbf{a}_{1,s,N}'\boldsymbol{\xi}_N \\ \frac{1}{2(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}[\mathbf{I}_T\otimes(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')]\mathbf{Q}_{0,N}\mathbf{v}_N+\mathbf{a}_{2,s,N}'\boldsymbol{\xi}_N \\ \boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N+\mathbf{a}_{3,s,N}'\boldsymbol{\xi}_N \\ \frac{1}{2}\boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}[\mathbf{I}_T\otimes(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')]\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N+\mathbf{a}_{4,s,N}'\boldsymbol{\xi}_N \\ \cdot \\ \frac{1}{(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}\mathbf{v}_N+\mathbf{a}_{a,N}'\boldsymbol{\xi}_N \\ \boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N+\mathbf{a}_{b,N}'\boldsymbol{\xi}_N \end{bmatrix}+o_p(1), \tag{25}$$

for $s=1,...,S$. The $(NT+N)\times 1$ vector $\boldsymbol{\xi}_N=(\mathbf{v}_N',\boldsymbol{\mu}_N')'$, $\mathbf{a}_{p,s,N}=T^{-1}\mathbf{T}_N\boldsymbol{\alpha}_{p,s,N}$, $p=1,...,4$, $s=1,...,S$, $\mathbf{a}_{a,N}=T^{-1}\mathbf{T}_N\boldsymbol{\alpha}_{a,N}$, and $\mathbf{a}_{b,N}=T^{-1}\mathbf{T}_N\boldsymbol{\alpha}_{b,N}$, which can also be written as

$\mathbf{a}_{p,s,N}=(\mathbf{a}_{p,s,N}^{v\,\prime},\mathbf{a}_{p,s,N}^{\mu\,\prime})'=T^{-1}[(\mathbf{T}_{v,N}\boldsymbol{\alpha}_{p,s,N})',(\mathbf{T}_{\mu,N}\boldsymbol{\alpha}_{p,s,N})']'$, $s=1,...,S$, $p=1,...,4$, and

$\mathbf{a}_{a,N}=(\mathbf{a}_{a,N}^{v\prime},\mathbf{a}_{a,N}^{\mu\prime})'=T^{-1}[(\mathbf{T}_{v,N}\boldsymbol{\alpha}_{a,N})',(\mathbf{T}_{\mu,N}\boldsymbol{\alpha}_{a,N})']'$,

$\mathbf{a}_{b,N}=(\mathbf{a}_{b,N}^{v\prime},\mathbf{a}_{b,N}^{\mu\prime})'=T^{-1}[(\mathbf{T}_{v,N}\boldsymbol{\alpha}_{b,N})',(\mathbf{T}_{\mu,N}\boldsymbol{\alpha}_{b,N})']'$.

Observe that the elements of $\mathbf{a}_{p,s,N}$, $p=1,...,4$, $s=1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$ are bounded uniformly in absolute value by Assumption 7 and Lemma B.1. Utilizing

(26)

$$\boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N=\mathbf{v}_N'\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{1,N}\mathbf{v}_N+T\boldsymbol{\mu}_N'\mathbf{M}_{s,N}'\mathbf{M}_{s,N}\boldsymbol{\mu}_N,$$

$$\frac{1}{2}\boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}[\mathbf{I}_T\otimes(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')]\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N=\frac{1}{2}\mathbf{v}_N'\mathbf{Q}_{1,N}[\mathbf{I}_T\otimes(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')]\mathbf{Q}_{1,N}\mathbf{v}_N+\frac{T}{2}\boldsymbol{\mu}_N'(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')\boldsymbol{\mu}_N,$$

$$\boldsymbol{\varepsilon}_N'\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N=\mathbf{v}_N'\mathbf{Q}_{1,N}\mathbf{v}_N+T\boldsymbol{\mu}_N'\boldsymbol{\mu}_N,$$

we have

(27)

$$N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N)=N^{-1/2}\begin{bmatrix}\frac{1}{(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{0,N}\mathbf{v}_N+\mathbf{a}_{1,s,N}'\boldsymbol{\xi}_N\\ \frac{1}{2(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}[\mathbf{I}_T\otimes(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')]\mathbf{Q}_{0,N}\mathbf{v}_N+\mathbf{a}_{2,s,N}'\boldsymbol{\xi}_N\\ \mathbf{v}_N'\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\mathbf{M}_{s,N}'\mathbf{M}_{s,N})\mathbf{Q}_{1,N}\mathbf{v}_N+T\boldsymbol{\mu}_N'\mathbf{M}_{s,N}'\mathbf{M}_{s,N}\boldsymbol{\mu}_N+\mathbf{a}_{3,s,N}'\boldsymbol{\xi}_N\\ \frac{1}{2}\mathbf{v}_N'\mathbf{Q}_{1,N}[\mathbf{I}_T\otimes(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')]\mathbf{Q}_{1,N}\mathbf{v}_N+\frac{T}{2}\boldsymbol{\mu}_N'(\mathbf{M}_{s,N}+\mathbf{M}_{s,N}')\boldsymbol{\mu}_N+\mathbf{a}_{4,s,N}'\boldsymbol{\xi}_N\\ \vdots\\ \frac{1}{(T-1)}\mathbf{v}_N'\mathbf{Q}_{0,N}\mathbf{v}_N+\mathbf{a}_{a,N}'\boldsymbol{\xi}_N\\ \mathbf{v}_N'\mathbf{Q}_{1,N}\mathbf{v}_N+T\boldsymbol{\mu}_N'\boldsymbol{\mu}_N+\mathbf{a}_{b,N}'\boldsymbol{\xi}_N\end{bmatrix}+o_p(1)$$

$$=N^{-1/2}\mathbf{q}_N^*+o_p(1)=\mathbf{q}_N+o_p(1).$$

Next, consider the $(4S+2)\times 1$ vector

$$\mathbf{q}_N=N^{-1/2}\mathbf{q}_N^*=N^{-1/2}\begin{bmatrix}\mathbf{q}_{1,N}^*\\ \vdots\\ \mathbf{q}_{S,N}^*\\ q_{a,N}^*\\ q_{b,N}^*\end{bmatrix}. \tag{28}$$

Each element $\mathbf{q}_{s,N}^*$, $s=1,...,S$, is a 4×1 vector, given by

$$\mathbf{q}_{s,N}^*=\begin{bmatrix}q_{1,s,N}^*\\ q_{2,s,N}^*\\ q_{3,s,N}^*\\ q_{4,s,N}^*\end{bmatrix}, \tag{29}$$

where $q^*_{p,s,N}$, $p=1,...,4$, $s=1,...,S$, $q^*_{a,N}$, and $q^*_{b,N}$ can be written as linear quadratic forms in the $(NT+N)\times 1$ vector $\xi_N = (\mathbf{v}'_N, \boldsymbol{\mu}'_N)'$, i.e., we have

$$q^*_{p,s,N} = \xi'_N \mathbf{A}_{p,s,N} \xi_N + \mathbf{a}'_{p,s,N} \xi_N, \; p=1,...,4, \; s=1,...,S, \tag{30}$$
$$q^*_{a,N} = \xi'_N \mathbf{A}_{a,N} \xi_N + \mathbf{a}'_{a,N} \xi_N, \text{ and}$$
$$q^*_{b,N} = \xi'_N \mathbf{A}_{b,N} \xi_N + \mathbf{a}'_{b,N} \xi_N.$$

We consider each of these terms in the following.

$$q^*_{1,s,N} = \xi'_N \mathbf{A}_{1,s,N} \xi_N + \mathbf{a}'_{1,s,N} \xi_N, \text{ where} \tag{31}$$
$$\mathbf{A}_{1,s,N} = \begin{bmatrix} \mathbf{A}^v_{1,s,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{A}^\mu_{1,s,N} \end{bmatrix} = \begin{bmatrix} \frac{1}{(T-1)} \mathbf{Q}_{0,N} (\mathbf{I}_T \otimes \mathbf{M}'_{s,N} \mathbf{M}_{s,N}) \mathbf{Q}_{0,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{0}_{N\times N} \end{bmatrix}, \text{ and}$$
$$\mathbf{a}'_{1,s,N} = (\mathbf{a}^{v\,\prime}_{1,s,N}, \mathbf{a}^{\mu\,\prime}_{1,s,N}),$$

and the $\mathbf{0}$ terms denote zero-matrices, whose dimensions are indicated by the subscript.

$$q^*_{2,s,N} = \xi'_N \mathbf{A}_{2,s,N} \xi_N + \mathbf{a}'_{2,s,N} \xi_N, \text{ where}$$
$$\mathbf{A}_{2,s,N} = \begin{bmatrix} \mathbf{A}^v_{2,s,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{A}^\mu_{2,s,N} \end{bmatrix} = \begin{bmatrix} \frac{1}{2(T-1)} \mathbf{Q}_{0,N} [\mathbf{I}_T \otimes (\mathbf{M}_{s,N} + \mathbf{M}'_{s,N})] \mathbf{Q}_{0,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{0}_{N\times N} \end{bmatrix}, \text{ and} \tag{32a}$$
$$\mathbf{a}'_{2,s,N} = (\mathbf{a}^{v\,\prime}_{2,s,N}, \mathbf{a}^{\mu\,\prime}_{2,s,N}).$$

$$q^*_{3,s,N} = \xi'_N \mathbf{A}_{3,s,N} \xi_N + \mathbf{a}'_{3,s,N} \xi_N, \text{ where}$$
$$\mathbf{A}_{3,s,N} = \begin{bmatrix} \mathbf{A}^v_{3,s,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{A}^\mu_{3,s,N} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{1,N} (\mathbf{I}_T \otimes \mathbf{M}'_{s,N} \mathbf{M}_{s,N}) \mathbf{Q}_{1,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & T\mathbf{M}'_{s,N} \mathbf{M}_{s,N} \end{bmatrix}, \text{ and} \tag{32b}$$
$$\mathbf{a}'_{3,s,N} = (\mathbf{a}^{v\,\prime}_{3,s,N}, \mathbf{a}^{\mu\,\prime}_{3,s,N}).$$

$$q^*_{4,s,N} = \xi'_N \mathbf{A}_{4,s,N} \xi_N + \mathbf{a}'_{4,s,N} \xi_N, \text{ where}$$
$$\mathbf{A}_{4,s,N} = \begin{bmatrix} \mathbf{A}^v_{4,s,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{A}^v_{4,s,N} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \mathbf{Q}_{1,N} [\mathbf{I}_T \otimes (\mathbf{M}_{s,N} + \mathbf{M}'_{s,N})] \mathbf{Q}_{1,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \frac{T}{2} (\mathbf{M}_{s,N} + \mathbf{M}'_{s,N}) \end{bmatrix}, \text{ and} \tag{32c}$$
$$\mathbf{a}'_{4,s,N} = (\mathbf{a}^{v\,\prime}_{4,s,N}, \mathbf{a}^{\mu\,\prime}_{4,s,N}).$$

$q^{*}_{a,N} = \boldsymbol{\xi}'_N \mathbf{A}_{a,N} \boldsymbol{\xi}_N + \mathbf{a}'_{a,N} \boldsymbol{\xi}_N$, where

$$\mathbf{A}_{a,N} = \begin{bmatrix} \mathbf{A}^{v}_{a,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{A}^{\mu}_{a,N} \end{bmatrix} = \begin{bmatrix} \frac{1}{(T-1)}\mathbf{Q}_{0,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{0}_{N\times N} \end{bmatrix}, \text{ and} \tag{32e}$$

$\mathbf{a}'_{a,N} = (\mathbf{a}^{v\prime}_{a,N}, \mathbf{a}^{\mu\prime}_{a,N})$.

$q^{*}_{b,N} = \boldsymbol{\xi}'_N \mathbf{A}_{b,N} \boldsymbol{\xi}_N + \mathbf{a}'_{b,N} \boldsymbol{\xi}_N$, where

$$\mathbf{A}_{b,N} = \begin{bmatrix} \mathbf{A}^{v}_{b,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \mathbf{A}^{\mu}_{b,N} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{1,N} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & T\mathbf{I}_N \end{bmatrix}, \text{ and} \tag{32f}$$

$\mathbf{a}'_{b,N} = (\mathbf{a}^{v\prime}_{b,N}, \mathbf{a}^{\mu\prime}_{b,N})$.

Note that the row and column sums of the symmetric $(NT+N)\times(NT+N)$ matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s=1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, are bounded uniformly in absolute value by Assumption 3 and Remark A.1. Moreover, the elements of the $\boldsymbol{\xi}_N = (\mathbf{v}'_N, \boldsymbol{\mu}'_N)'$ are independently distributed by Assumption 1. Hence, the variance-covariance matrix of $\boldsymbol{\xi}_N$ is

$$\boldsymbol{\Omega}_{\xi,N} = \begin{bmatrix} \sigma^2_v \mathbf{I}_{NT} & \mathbf{0}_{NT\times N} \\ \mathbf{0}_{N\times NT} & \sigma^2_\mu \mathbf{I}_N \end{bmatrix}. \tag{33}$$

In order to calculate the variance-covariance matrix of $\mathbf{q}_N$, denoted as $\boldsymbol{\Psi}_N$, we invoke Lemma A.1 in Kelejian and Prucha (2008). It is given by $\boldsymbol{\Psi}_N = N^{-1}E(\mathbf{q}^{*}_N \mathbf{q}^{*\prime}_N)$, which is a symmetric $(4S+2)\times(4S+2)$ matrix, and takes the following form:

$$\boldsymbol{\Psi}_N = (\mathfrak{E}_{r,s,N}),\ r,s=1,...,S+1, \text{ i.e.,} \tag{34a}$$

$$\boldsymbol{\Psi}_N = \begin{bmatrix} \mathfrak{E}_{1,1,N} & \cdot & \mathfrak{E}_{1,S,N} & \mathfrak{E}_{1,S+1,N} \\ \cdot & & & \cdot \\ \mathfrak{E}_{S,1,N} & & \mathfrak{E}_{S,S,N} & \mathfrak{E}_{S,S+1,N} \\ \mathfrak{E}_{S+1,1,N} & & \mathfrak{E}_{S+1,S,N} & \mathfrak{E}_{S+1,S+1,N} \end{bmatrix}. \tag{34b}$$

Observe that the matrix $\boldsymbol{\Psi}_N$ contains three parts.

i) The upper left block is of dimension $4S\times 4S$, consisting of S^2 blocks of dimension 4 × 4, which are defined as

$$\mathfrak{E}_{r,s,N} = N^{-1}E(\mathbf{q}^*_{r,N}\mathbf{q}^{*\prime}_{s,N}) = (\mathfrak{E}^{p,q}_{r,s}),\ r,s=1,...,S,\ p,q=1,...,4. \tag{34c}$$

The elements $\mathfrak{E}^{p,q}_{r,s,N}$, $p,q=1,...,4$, $r,s=1,...,S$ are defined as

$$\begin{aligned}
&\mathfrak{E}^{p,q}_{r,s,N} = N^{-1}Cov(q^*_{p,r,N}, q^*_{q,s,N}) \qquad (34d)\\
&= 2\sigma_v^4 N^{-1}Tr(\mathbf{A}^v_{p,r,N}\mathbf{A}^v_{q,s,N}) + 2\sigma_\mu^4 N^{-1}Tr(\mathbf{A}^\mu_{p,r,N}\mathbf{A}^\mu_{q,s,N})\\
&+ \sigma_v^2 N^{-1}\mathbf{a}^{v\prime}_{p,r,N}\mathbf{a}^v_{q,s,N} + \sigma^2_{\mu,N}N^{-1}\mathbf{a}^{\mu\prime}_{p,r,N}\mathbf{a}^\mu_{q,s,N}\\
&+ (\sigma_v^{(4)} - 3\sigma_v^4)N^{-1}\sum_{i=1}^{NT} a^v_{p,r,ii,N}a^v_{q,s,ii,N} + (\sigma_\mu^{(4)} - 3\sigma_\mu^4)N^{-1}\sum_{i=1}^{N} a^\mu_{p,r,ii,N}a^\mu_{q,s,ii,N}\\
&+ \sigma_v^{(3)}N^{-1}\sum_{i=1}^{NT}(a^v_{p,r,i,N}a^v_{q,s,ii,N} + a^v_{p,r,ii,N}a^v_{q,s,i,N}) + \sigma_\mu^{(3)}N^{-1}\sum_{i=1}^{N}(a^\mu_{p,r,i,N}a^\mu_{q,s,ii,N} + a^\mu_{p,r,ii,N}a^\mu_{q,s,i,N}),
\end{aligned}$$

where $a^v_{p,r,ii,N}$ and $a^\mu_{p,r,ii,N}$ denote the *i*-th main diagonal element of the matrices $\mathbf{A}^v_{p,r,N}$ and $\mathbf{A}^\mu_{p,r,N}$, respectively, and $a^v_{p,r,i,N}$ and $a^v_{p,r,i,N}$ denote the *i*-th element of the vectors $\mathbf{a}^v_{p,r,N}$ and $\mathbf{a}^\mu_{p,r,N}$ respectively. The terms $\sigma_v^{(3)}$, $\sigma_\mu^{(3)}$ and $\sigma_v^{(4)}$, $\sigma_\mu^{(4)}$ denote the third and fourth moment of $v_{it,N}$ and $\mu_{it,N}$, respectively.

ii) The last two rows and columns are matrices of dimension $(2\times 4S)$ and $(4S\times 2)$, respectively, each of which is made up by S blocks of dimension (2×4) (4×2), defined as

$$\mathfrak{E}_{S+1,s,N} = (\mathfrak{E}^{p,q}_{S+1,s,N}) = N^{-1}Cov(q^*_{p,N}, q^*_{q,s,N}),\ p=a,b,\ q=1,...,4, \text{ and } s=1,...,S, \tag{34e}$$

and $\mathfrak{E}_{s,S+1} = (\mathfrak{E}_{S+1,s})'$, $s=1,...,S$. The elements $\mathfrak{E}^{p,q}_{S+1,s,N}$ are defined as in (34d), using the corresponding indexation.

iii) Finally, the lower right block of dimension 2 × 2, is defined as

$$\mathfrak{E}_{S+1,S+1,N} = (\mathfrak{E}^{p,q}_{S+1,S+1,N}) = N^{-1}Cov(q^*_{p,N}, q^*_{q,N}),\ p,q=a,b, \tag{34f}$$

where the elements $\mathfrak{E}^{p,q}_{S+1,s,N}$ are defined as in (34d), using the corresponding indexation.

For definiteness, we add that the position of each block $\mathfrak{E}_{r,s,N}$ is such that its upper left element appears in row $(4r-3)$ and column $(4s-3)$ of the $(4S+2)\times(4S+2)$ matrix $\mathbf{\Psi}_N$. The position of each block $\mathfrak{E}_{S+1,s}$, $s=1,...,S$, is such that its first element appears in row

$(4S+1)$ and column $(4s-3)$ of $\boldsymbol{\Psi}_N$. Finally, the upper left element of the block ($\mathfrak{E}_{S+1,S+1,N}$) appears in row $(4S+1)$ and column $(4S+1)$ of $\boldsymbol{\Psi}_N$.

The expression given by (34d) holds generally. Part of the elements of $\boldsymbol{\Psi}_N$ can be stated in simpler terms, considering the matrices and vectors in (32), which appear in the quadratic forms. In particular, the submatrices $\mathbf{A}^{\mu}_{p,s,N}$, $p=1,2$, and $\mathbf{A}^{\mu}_{a,N}$ are zero such that the second term in the first line of (34d), and the second term in the third line, drop out for $p=1,2$ or $q=1,2$ and where $\mathbf{A}_{a,N}$ is involved. If both sub-matrices associated with $\mu_{it,N}$ are zero ($p=q=1,2$, or where $\mathbf{A}_{a,N}$ appears twice), the second term in the fourth line drops out as well. Further, the matrices $\mathbf{A}_{p,s,N}$ have zero main diagonal elements for $p=2,4$ such that the terms involving the fourth moments are zero for $p=2,4$ or $q=2,4$; for $p=2,4$ and $q=2,4$ or $q=4$, the expressions involving the third moments of $v_{it,N}$ and $\mu_{it,N}$ are zero as well. Moreover, due to the orthogonality of $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$, the terms in the first line drop out when $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ meet in the trace expression. Finally, if $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ are normally distributed, the terms involving the third and fourth moments of $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ drop out for all elements of $\boldsymbol{\Psi}_N$; we denote the variance-covariance matrix under the assumption that $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ are normally distributed as $\boldsymbol{\Psi}^{\circ}_N$. If $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ are not normally distributed, $\boldsymbol{\Psi}^{\circ}_N$ can be regarded as an approximation of the true matrix $\boldsymbol{\Psi}_N$.

To derive the asymptotic distribution of $\mathbf{q}_N$ and $\widetilde{\boldsymbol{\theta}}_N$ we invoke the central limit theorem for vectors of linear quadratic forms given by Kelejian and Prucha (2008, Theorem A.1) and Corollary F4 in Pötscher and Prucha (1997). We summarize the results regarding the asymptotic distribution of $\widetilde{\boldsymbol{\theta}}_N$ in the following Theorem, which is proved in Appendix B.

Theorem 2. (Asymptotic Normality of $\widetilde{\boldsymbol{\theta}}_N$)

Let $\widetilde{\boldsymbol{\theta}}_N$ be the GM estimator defined by (18). Suppose Assumptions 1-7 hold and, furthermore, that $\lambda_{\min}(\boldsymbol{\Psi}_N) \geq c^*_{\boldsymbol{\Psi}} > 0$. Then, provided the optimization space contains the parameter space, we have

$$N^{1/2}(\widetilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = (\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N^{1/2}\boldsymbol{\xi}_N + o_p(1), \text{ with}$$

$$\mathbf{J}_N = \frac{\partial}{\partial\boldsymbol{\theta}'}\boldsymbol{\Gamma}_N\mathfrak{b}_N = \boldsymbol{\Gamma}_N\mathfrak{B}_N, \text{ and}$$

$$\boldsymbol{\xi}_N = \boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} N(0, \mathbf{I}_{4S+2}),$$

where $\boldsymbol{\Psi}_N = E(\mathbf{q}_N \mathbf{q}_N')$ and $\boldsymbol{\Psi}_N = (\boldsymbol{\Psi}_N^{1/2})(\boldsymbol{\Psi}_N^{1/2})'$.

Furthermore $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = O_p(1)$ and

$$\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\boldsymbol{\Theta}_N) = (\mathbf{J}_N'\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}_N'\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N\boldsymbol{\Theta}_N\mathbf{J}_N(\mathbf{J}_N'\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1},$$

where $\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}$ is positive definite.

Theorem 2 implies that the difference between the cumulative distribution function of $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ and that of $N(0, \boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N})$ converges pointwise to zero, which justifies the use of the latter as an approximation of the former.[9]

Note that $\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\boldsymbol{\Psi}_N^{-1}) = (\mathbf{J}_N'\boldsymbol{\Psi}_N^{-1}\mathbf{J}_N)^{-1}$ and that $\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\boldsymbol{\Theta}_N) - \boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\boldsymbol{\Psi}_N^{-1})$ is positive semidefinite. Thus, using a consistent estimator of $\boldsymbol{\Psi}_N^{-1}$ (which will be derived below) as weighting matrix $\boldsymbol{\Theta}_N$ leads to the efficient GM estimator. We add that $\boldsymbol{\Psi}_N$ is not exactly equal to the variance-covariance matrix of the moments, since the GM estimates are based on estimated rather than the true disturbances and since there is an endogenous right-hand side variable in equation (1).

Remark 1. Under normality, Theorem 2 holds, with $\boldsymbol{\Psi}_N$ replaced by $\boldsymbol{\Psi}_N^{\circ}$. If $\boldsymbol{\varepsilon}_N$ is not normal the use of $\boldsymbol{\Psi}_N^{\circ}$ delivers an approximation to the true variance-covariance matrix of $\tilde{\boldsymbol{\theta}}_N$.

3.3 Estimation of the Variance-Covariance Matrix of $\tilde{\boldsymbol{\theta}}_N$

In the following, we develop a consistent estimator for the variance-covariance matrix of $\tilde{\boldsymbol{\theta}}_N$. Define

$$\tilde{\mathbf{J}}_N = \tilde{\boldsymbol{\Gamma}}_N\tilde{\mathfrak{B}}_N. \tag{35}$$

We next specify estimators for $\mathbf{a}_{p,s,N} = \mathbf{T}_N\boldsymbol{\alpha}_{p,s,N}$, $p = 1,...,4$, $s = 1,...,S$, $\mathbf{a}_{a,N} = \mathbf{T}_N\boldsymbol{\alpha}_{a,N}$, and $\mathbf{a}_{b,N} = \mathbf{T}_N\boldsymbol{\alpha}_{b,N}$. The matrix $\mathbf{T}_N$ will often be of the form

$$\mathbf{T}_N = \mathbf{F}_N\mathbf{P}_N \quad \text{with } \mathbf{F}_N = (\mathbf{F}_{v,N}', \mathbf{F}_{\mu,N}')', \tag{36a}$$

[9] Compare Corollary F4 in Pötscher and Prucha (1997).

which can also be written as

$$\mathbf{T}_N = (\mathbf{T}'_{v,N}, \mathbf{T}'_{\mu,N})' \text{ with } \mathbf{T}_{v,N} = \mathbf{F}_{v,N}\mathbf{P}_N,\ \mathbf{T}_{\mu,N} = \mathbf{F}_{\mu,N}\mathbf{P}_N,$$

and

$$\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N, \tag{36b}$$

$$\mathbf{F}_{\mu,N} = (\mathbf{e}'_T \otimes \mathbf{I}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N,$$

or, alternatively,

$$\mathbf{F}_{v,N} = \mathbf{\Omega}_{\varepsilon,N}^{-1}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N, \tag{36c}$$

$$\mathbf{F}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N,$$

where $\mathbf{F}_{v,N}$ is a real nonstochastic $NT \times P_*$ matrix, $\mathbf{F}_{\mu,N}$ is a real nonstochastic $N \times P_*$ matrix, $\mathbf{H}_N$ is a real nonstochastic $NT \times P_*$ matrix of instruments, and $\mathbf{P}_N$ is a real nonstochastic $P_* \times P$ matrix, with P as in Assumption 7.

To be more specific, when equation (1a) is estimated using twp-stages least squares (TSLS), $\mathbf{\Delta}_N = (\tilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N)$ and the matrix $\mathbf{P}_N$ will be of the structure as defined above and can be estimated consistently by some estimator $\tilde{\mathbf{P}}_N$ (see Section IV).

The estimators for $\mathbf{T}_N$ are defined as

$$\tilde{\mathbf{T}}_{v,N} = \tilde{\mathbf{F}}_{v,N}\tilde{\mathbf{P}}_N,\ \ \tilde{\mathbf{T}}_{\mu,N} = \tilde{\mathbf{F}}_{\mu,N}\tilde{\mathbf{P}}_N, \tag{37a}$$

$$\tilde{\mathbf{F}}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \tilde{\rho}_{m,N}\mathbf{M}'_{m,N})^{+}]\mathbf{H}_N, \text{ or} \tag{37b}$$

$$\tilde{\mathbf{F}}_{\mu,N} = (\mathbf{e}'_T \otimes \mathbf{I}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \tilde{\rho}_{m,N}\mathbf{M}'_{m,N})^{+}]\mathbf{H}_N,$$

or

$$\widetilde{\mathbf{F}}_{v,N} = \widetilde{\mathbf{\Omega}}_{\boldsymbol{\varepsilon},N}^{-1}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \widetilde{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N\,, \tag{37c}$$

$$\widetilde{\mathbf{F}}_{\mu,N} = [\widetilde{\sigma}_{1,N}^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \widetilde{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N\,.$$

The estimators of $\mathbf{a}_{p,s,N} = \mathbf{T}_N\boldsymbol{\alpha}_{p,s,N}$, $p = 1,...,4,a,b$, $s = 1...,S$, $\mathbf{a}_{a,N} = \mathbf{T}_N\boldsymbol{\alpha}_{a,N}$, and $\mathbf{a}_{b,N} = \mathbf{T}_N\boldsymbol{\alpha}_{b,N}$ are then given by

$$\widetilde{\mathbf{a}}_{p,s,N} = \widetilde{\mathbf{T}}_N\widetilde{\boldsymbol{\alpha}}_{p,s,N} \tag{38}$$

with $\widetilde{\boldsymbol{\alpha}}_{p,s,N} = 2N^{-1}(\mathbf{D}_N'\widetilde{\mathbf{C}}_{p,s,N}\widetilde{\mathbf{u}}_N)$, and the matrices $\widetilde{\mathbf{C}}_{p,s,N}$, $p = 1,...,4$, $s = 1,...,S$, $\widetilde{\mathbf{C}}_{a,N}$, and $\widetilde{\mathbf{C}}_{b,N}$ are given by (22) with $\boldsymbol{\rho}_N$ replaced by $\widetilde{\boldsymbol{\rho}}_N$.

The elements of the estimated $(4S+2)\times(4S+2)$ matrix $\widetilde{\mathbf{\Psi}}_N$ are defined in (34d), with $\sigma_{v,N}$ and $\sigma_{\mu,N}$ replaced by $\widetilde{\sigma}_{v,N}$ and $\widetilde{\sigma}_{v,N}$. The third and fourth moments of $\mu_{i,N}$ and $v_{it,N}$, denoted as $\sigma_\mu^{(3)},\sigma_v^{(3)}$ and $\sigma_\mu^{(4)},\sigma_v^{(4)}$, can be estimated consistently as follows (see Appendix B):

$$\widetilde{\sigma}_{\varepsilon,N}^{(3)} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\varepsilon}_{it,N}^3\,, \tag{39a}$$

$$\widetilde{\sigma}_{\mu,N}^{(3)} = \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it}^2\,, \tag{39b}$$

$$\widetilde{\sigma}_{v,N}^{(3)} = \widetilde{\sigma}_{\varepsilon,N}^{(3)} - \widetilde{\sigma}_{\mu,N}^{(3)}\,, \tag{39c}$$

as well as

$$\widetilde{\sigma}_{\mu,N}^{(4)} = \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it,N}^3 \tag{40a}$$

$$-\frac{3}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it,N}(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\varepsilon}_{it,N}^2 - \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it,N})\,, \tag{40b}$$

$$\widetilde{\sigma}_{v,N}^{(4)} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\varepsilon}_{it,N}^4 - \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it,N}^3 \tag{40c}$$

$$-\frac{3}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it,N}(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\varepsilon}_{it,N}^2 - \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1\\t\neq s}}^{T}\widetilde{\varepsilon}_{is,N}\widetilde{\varepsilon}_{it,N})\,,$$

where $\tilde{\boldsymbol{\varepsilon}}_N = \mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \tilde{\rho}_{m,N}\mathbf{M}_{m,N})\tilde{\mathbf{u}}_N$.[10] Based on $\tilde{\boldsymbol{\Psi}}_N$, we can now define the estimator for $\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}$ as

$$\tilde{\boldsymbol{\Omega}}_{\tilde{\boldsymbol{\theta}}_N}(\tilde{\boldsymbol{\Theta}}_N) = (\tilde{\mathbf{J}}_N'\tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N)^{+}\tilde{\mathbf{J}}_N'\tilde{\boldsymbol{\Theta}}_N\tilde{\boldsymbol{\Psi}}_N\tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N(\tilde{\mathbf{J}}_N'\tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N)^{+}. \tag{41}$$

The following theorem establishes the consistency of $\tilde{\boldsymbol{\Psi}}_N$ and $\tilde{\boldsymbol{\Omega}}_{\tilde{\boldsymbol{\theta}}_N}$.

Theorem 3. Variance-Covariance Matrix Estimation

Suppose all of the assumptions of Theorem 2, apart from Assumption 5, hold and that additionally all of the fourth moments of the elements of $\mathbf{D}_N$ are bounded uniformly. Suppose furthermore (a) that the elements of the nonstochastic matrices $\mathbf{H}_N$ are bounded uniformly in absolute value, (b) $\sup_N \sum_{s=1}^{S} |\rho_{s,N}| < 1$ and that the row and column sums of $\mathbf{M}_N$ are bounded uniformly in absolute value by one and some finite constant respectively, and (c) $\tilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$ with $\mathbf{P}_N = O(1)$. Then, $\tilde{\boldsymbol{\Psi}}_N - \boldsymbol{\Psi}_N = o_p(1)$ and $\tilde{\boldsymbol{\Psi}}_N^{-1} - \boldsymbol{\Psi}_N^{-1} = o_p(1)$. Furthermore, if Assumption 5 holds, then also $\tilde{\boldsymbol{\Omega}}_{\tilde{\boldsymbol{\theta}}_N} - \boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N} = o_p(1)$.

Remark 1.

Theorem 3 also holds, if $\tilde{\boldsymbol{\theta}}_N$ is replaced by some other estimator $(NT)^{1/2}(\breve{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = O_p(1)$. Notice that condition (b) can be dropped in case that $\mathbf{F}_{v,N} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$ and $\mathbf{F}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$. The consistency result for $\tilde{\boldsymbol{\Psi}}_N^{-1}$ verifies that this estimator for $\boldsymbol{\Psi}_N^{-1}$ can indeed be used in the formulation of an efficient GM estimator.

3. Joint Distribution of the GM Estimator for $\boldsymbol{\theta}_N$ and Estimators of Other Model Parameters

Note that both $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ and $(NT)^{1/2}\boldsymbol{\Delta}_N$, and thus also $N^{1/2}\boldsymbol{\Delta}_N$ are asymptotically linear in $\boldsymbol{\xi}_N$. Hence, the joint distribution of the vector $[N^{1/2}\boldsymbol{\Delta}_N', N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)']'$ can be derived invoking the central limit theorem for vectors of quadratic forms by Kelejian and Prucha (2008).

[10] Compare Gilbert (2002) for the estimation of third and fourth moments in error component models without spatial lags and without spatial autoregressive disturbances.

Consider the $(P_* + 4S + 2)\times 1$ vector of linear and linear quadratic forms in $\boldsymbol{\xi}_N$:

$$\mathbf{w}_N = \begin{bmatrix} (NT)^{-1/2}\mathbf{F}_N'\boldsymbol{\xi}_N \\ \mathbf{q}_N \end{bmatrix}. \tag{42}$$

Its variance-covariance matrix is of dimension $(P_* + 4S + 2)\times(P_* + 4S + 2)$ and given by:

$$Var(\mathbf{w}_N) = \boldsymbol{\Psi}_{o,N} = E\begin{bmatrix} (NT)^{-1}\mathbf{F}_N'\boldsymbol{\xi}_N\boldsymbol{\xi}_N'\mathbf{F}_N & (NT)^{-1/2}\mathbf{F}_N'\boldsymbol{\xi}_N\mathbf{q}_N' \\ (NT)^{-1/2}\mathbf{q}_N\boldsymbol{\xi}_N'\mathbf{F}_N & \mathbf{q}_N\mathbf{q}_N' \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}_{\Delta\Delta,N} & \boldsymbol{\Psi}_{\Delta\theta,N} \\ \boldsymbol{\Psi}_{\Delta\theta,N}' & \boldsymbol{\Psi}_N \end{bmatrix}, \tag{43a}$$

where the $(4S + 2)\times(4S + 2)$ matrix $\boldsymbol{\Psi}_N$ is defined above, $\boldsymbol{\Psi}_{\Delta\Delta,N}$ is of dimension $P_*\times P_*$ and defined as

$$\boldsymbol{\Psi}_{\Delta\Delta,N} = E[(NT)^{-1}\mathbf{F}_N'\boldsymbol{\xi}_N\boldsymbol{\xi}_N'\mathbf{F}_N] = (NT)^{-1}(\sigma_v^2\mathbf{F}_{v,N}'\mathbf{F}_{v,N} + \sigma_\mu^2\mathbf{F}_{\mu,N}'\mathbf{F}_{\mu,N}), \tag{43b}$$

and the $P_*\times(4S + 2)$ matrix $\boldsymbol{\Psi}_{\Delta\theta,N}$ is given by

$$\begin{aligned} &\boldsymbol{\Psi}_{\Delta\theta,N} = E[(NT)^{-1/2}\mathbf{F}_N'\boldsymbol{\xi}_N\mathbf{q}_N'] \\ &= (NT)^{-1/2}N^{-1/2}[\mathbf{F}_{v,N}'(\sigma_v^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{1,1,N}^v} + \sigma_v^2\mathbf{a}_{1,1,N}^v) + \mathbf{F}_{\mu,N}'(\sigma_\mu^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{1,1,N}^\mu} + \sigma_\mu^2\mathbf{a}_{1,1,N}^\mu),..., \\ &...,\mathbf{F}_{v,N}'(\sigma_v^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{a,N}^v} + \sigma_v^2\mathbf{a}_{a,N}^v) + \mathbf{F}_{\mu,N}'(\sigma_\mu^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{a,N}^\mu} + \sigma_\mu^2\mathbf{a}_{a,N}^\mu), \\ &\mathbf{F}_{v,N}'(\sigma_v^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{b,N}^v} + \sigma_v^2\mathbf{a}_{b,N}^v) + \mathbf{F}_{\mu,N}'(\sigma_\mu^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{b,N}^\mu} + \sigma_\mu^2\mathbf{a}_{b,N}^\mu), \end{aligned} \tag{43c}$$

where $\boldsymbol{\kappa}_{\mathbf{A}_{p,q,N}^v}$ and $\boldsymbol{\kappa}_{\mathbf{A}_{p,q,N}^\mu}$ are $NT\times 1$ and $N\times 1$ vectors, whose *i*-th element corresponds to the *i*-th main diagonal element of $\mathbf{A}_{p,q,N}^v$ and $\mathbf{A}_{p,q,N}^\mu$, respectively.

As we demonstrate in Appendix B, the matrix $\boldsymbol{\Psi}_{o,N}$ can be estimated consistently by

$$\tilde{\boldsymbol{\Psi}}_{o,N} = \begin{bmatrix} \tilde{\boldsymbol{\Psi}}_{\Delta\Delta,N} & \tilde{\boldsymbol{\Psi}}_{\Delta\theta,N} \\ \tilde{\boldsymbol{\Psi}}_{\Delta\theta,N}' & \tilde{\boldsymbol{\Psi}}_N \end{bmatrix}, \text{ where} \tag{44}$$

$$\begin{aligned} &\tilde{\boldsymbol{\Psi}}_{\Delta\Delta,N} = (NT)^{-1}[\tilde{\sigma}_v^2\tilde{\mathbf{F}}_{v,N}'\tilde{\mathbf{F}}_{v,N} + \tilde{\sigma}_\mu^2\tilde{\mathbf{F}}_{\mu,N}'\tilde{\mathbf{F}}_{\mu,N}], \\ &\tilde{\boldsymbol{\Psi}}_{\Delta\theta,N} = (NT)^{-1/2}N^{-1/2}[\tilde{\mathbf{F}}_{v,N}'(\tilde{\sigma}_v^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{1,1,N}^v} + \tilde{\sigma}_v^2\tilde{\mathbf{a}}_{1,1,N}^v) + \tilde{\mathbf{F}}_{\mu,N}'(\tilde{\sigma}_\mu^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{1,1,N}^\mu} + \tilde{\sigma}_\mu^2\tilde{\mathbf{a}}_{1,1,N}^\mu),..., \\ &...,\tilde{\mathbf{F}}_{v,N}'(\tilde{\sigma}_v^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{a,N}^v} + \tilde{\sigma}_v^2\tilde{\mathbf{a}}_{a,N}^v) + \tilde{\mathbf{F}}_{\mu,N}'(\tilde{\sigma}_\mu^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{a,N}^\mu} + \tilde{\sigma}_\mu^2\tilde{\mathbf{a}}_{a,N}^\mu), \\ &\tilde{\mathbf{F}}_{v,N}'(\tilde{\sigma}_v^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{b,N}^v} + \tilde{\sigma}_v^2\tilde{\mathbf{a}}_{b,N}^v) + \tilde{\mathbf{F}}_{\mu,N}'(\tilde{\sigma}_\mu^{(3)}\boldsymbol{\kappa}_{\mathbf{A}_{b,N}^\mu} + \tilde{\sigma}_\mu^2\tilde{\mathbf{a}}_{b,N}^\mu)]. \end{aligned}$$

Regarding the joint limiting distribution of $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ and $(NT)^{1/2}\boldsymbol{\Delta}_N$, we now have the following result.

Theorem 4. Joint Distribution of $\tilde{\boldsymbol{\theta}}_N$ and Other Model Parameters

Suppose all assumptions used in Theorem 3 hold and $\lambda_{\min}(\boldsymbol{\Psi}_{o,N}) \geq c^*_{\boldsymbol{\Psi}_o} > 0$. Then,

$$\begin{bmatrix} N^{1/2}\boldsymbol{\Delta}_N \\ N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) \end{bmatrix} = \begin{bmatrix} T^{-1/2}\mathbf{P}'_N & \mathbf{0} \\ \mathbf{0} & (\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N \end{bmatrix} \boldsymbol{\Psi}_{o,N}^{1/2}\boldsymbol{\xi}_{o,N} + o_p(1), \text{ with}$$

$$\boldsymbol{\xi}_{o,N} = \boldsymbol{\Psi}_{o,N}^{-1/2}[N^{-1/2}\boldsymbol{\xi}'_N\mathbf{F}_N, \mathbf{q}'_N]' \xrightarrow{d} N(\mathbf{0}, \mathbf{I}_{P^*+4S+2}),$$

$$\boldsymbol{\Omega}_{o,N} = \begin{bmatrix} T^{-1/2}\mathbf{P}'_N & \mathbf{0} \\ \mathbf{0} & (\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N \end{bmatrix} \boldsymbol{\Psi}_{o,N} \begin{bmatrix} T^{-1/2}\mathbf{P}_N & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Theta}_N\mathbf{J}_N(\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1} \end{bmatrix}, \text{ and}$$

$$\tilde{\boldsymbol{\Omega}}_{o,N} = \begin{bmatrix} T^{-1/2}\tilde{\mathbf{P}}'_N & \mathbf{0} \\ \mathbf{0} & (\tilde{\mathbf{J}}'_N\tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N)^{+}\tilde{\mathbf{J}}'_N\tilde{\boldsymbol{\Theta}}_N \end{bmatrix} \tilde{\boldsymbol{\Psi}}_{o,N} \begin{bmatrix} T^{-1/2}\tilde{\mathbf{P}}_N & \mathbf{0} \\ \mathbf{0} & \tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N(\tilde{\mathbf{J}}'_N\tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N)^{+} \end{bmatrix}.$$

Moreover,

$$\tilde{\boldsymbol{\Psi}}_{o,N} - \boldsymbol{\Psi}_{o,N} = o_p(1),\ \tilde{\boldsymbol{\Omega}}_{o,N} - \boldsymbol{\Omega}_{o,N} = o_p(1), \text{ and } \boldsymbol{\Psi}_{o,N} = O(1),\ \boldsymbol{\Omega}_{o,N} = O(1).$$

Theorem 4 implies that the difference between the joint cumulative distribution function of $[N^{1/2}\boldsymbol{\Delta}'_N, N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)']'$ and that of $N(\mathbf{0}, \boldsymbol{\Omega}_{o,N})$ converges pointwise to zero, which justifies the use of the latter distribution as an approximation of the former. The theorem also states that $\tilde{\boldsymbol{\Omega}}_{o,N}$ is a consistent estimator of $\boldsymbol{\Omega}_{o,N}$. The proof of Theorem 4 is given in Appendix B.

Remark 2.

As in Kelejian and Prucha (2008, p. 17), Theorem 4 can also be used to obtain the joint distribution of $(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ and some other estimator $\boldsymbol{\Delta}^{**}_N$, where $(NT)^{1/2}\boldsymbol{\Delta}^{**}_N = (NT)^{1/2}\mathbf{T}^{**\prime}_N\boldsymbol{\xi}_N + o_p(1)$, $\mathbf{T}^{**}_N = \mathbf{F}^{**}_N\mathbf{P}^{**}_N$, $\tilde{\mathbf{T}}^{**}_N = \tilde{\mathbf{F}}^{**}_N\tilde{\mathbf{P}}^{**}_N$, assuming that analogous assumptions are maintained for this estimator. In particular, the results remain valid, but with $\mathbf{F}_N$, $\mathbf{P}_N$ replaced by $\mathbf{F}^{**}_N$, $\mathbf{P}^{**}_N$, and $\tilde{\mathbf{F}}_N$, $\tilde{\mathbf{P}}_N$ replaced by $\tilde{\mathbf{F}}^{**}_N$, $\tilde{\mathbf{P}}^{**}_N$, in the definitions of $\boldsymbol{\Psi}_{\Delta\Delta,N}$, $\boldsymbol{\Psi}_{\Delta\theta,N}$, $\tilde{\boldsymbol{\Psi}}_{\Delta\Delta,N}$, and $\tilde{\boldsymbol{\Psi}}_{\Delta\theta,N}$.

IV. Two-Stages Least Squares (TSLS) Estimator for $\boldsymbol{\delta}_N$

1. Instruments

It is evident from model (1), that $E(\overline{\mathbf{Y}}_N \mathbf{u}'_N) \neq \mathbf{0}$. In line with Kelejian and Prucha (2008), we consider a TSLS procedure to obtain consistent estimates of the parameters $\boldsymbol{\delta}_N$. The following assumptions are maintained.

Assumption 8.

The regressor matrix $\mathbf{X}_N$ has full column rank (for N large enough). Furthermore, the elements of $\mathbf{X}_N$ are bounded uniformly in absolute value.

Assumption 9.

The instrument matrix $\mathbf{H}_N$ has full column rank $P_* \geq K + R$ (for N large enough). Furthermore, the elements of $\mathbf{H}_N$ are bounded uniformly in absolute value.

Assumption 10.

$\mathbf{Q}_{\mathbf{HH}} = \lim_{N\to\infty}[(NT)^{-1}\mathbf{H}'_N\mathbf{H}_N]$ is finite and nonsingular.

$\mathbf{Q}_{\mathbf{HZ}} = \operatorname{plim}_{N\to\infty}[(NT)^{-1}\mathbf{H}'_N\mathbf{Z}_N]$ is finite and non-singular.

Regarding the choice of instruments, note that

$$E(\sum_{r=1}^{R}\mathbf{W}_{r,N}\mathbf{y}_N) = \sum_{r=1}^{R}\mathbf{W}_{r,N}E(\mathbf{y}_N) = \sum_{r=1}^{R}\mathbf{W}_{r,N}E\{[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{r=1}^{R}\lambda_{r,N}\mathbf{W}_{r',N})^{-1}]\mathbf{X}_N\boldsymbol{\beta}_N\}$$

$$= \sum_{r=1}^{R}\mathbf{W}_{r,N}\{\mathbf{I}_T \otimes [\mathbf{I}_N + \sum_{i=1}^{\infty}(\sum_{r'=1}^{R}\lambda_{r',N}\mathbf{W}_{r',N})^i]\}\mathbf{X}_N\boldsymbol{\beta}_N, \qquad (45)$$

provided that $\left\|\sum_{r'=1}^{R}\lambda_{r',N}\mathbf{W}_{r',N}\right\| < 1$ for some matrix norm $\|\cdot\|$ (compare Horn and Johnson, 1985, p. 301). The instrument matrices $\mathbf{H}_N$ are used to instrument $\mathbf{Z}_N = (\mathbf{X}_N, \overline{\mathbf{Y}}_N)$ in terms of their predicted values from a least squares regression on $\mathbf{H}_N$, i.e., $\hat{\mathbf{Z}}_N = \mathbf{P}_{\mathbf{H}_N}\mathbf{Z}_N$, where $\mathbf{P}_{\mathbf{H}_N} = \mathbf{H}_N(\mathbf{H}'_N\mathbf{H}_N)^{-1}\mathbf{H}'_N$. In light of (45) it is reasonable to select $\mathbf{H}_N$ to include $\mathbf{X}_N$ and a subset of the linearly independent columns of terms of the sum

$$[\mathbf{I}_T \otimes \sum_{i=1}^{Q}(\sum_{r'=1}^{R}\mathbf{W}_{r',N})^i]\mathbf{X}_N, \qquad (46)$$

where Q is some predefined constant.[11] Note that such a choice of $\mathbf{H}_N$ implies that the second part of Assumption 9 will be fulfilled (by Assumptions 3 and 8) and that $\mathbf{X}_N$ is instrumented by itself.

2. Definition of TSLS Estimator and Asymptotic Results

Estimation of the model in equation (1) proceeds in three steps. In the first step, model (1a) is estimated by TSLS using instruments $\mathbf{H}_N$. In the second step, the spatial autoregressive parameters $\rho_{1,N},...,\rho_{S,N}$, and the variances of the error components σ_v^2 and σ_1^2 are estimated using the GM estimators defined in Section III in (17) and (18), based on consistent estimates of $\mathbf{u}_N$ from the first step. In the third step, the model is re-estimated by feasible generalized two-stages least squares (FGTSLS), which is equivalent to performing a TSLS estimation on a transformed version of equation (1). The advantage of this approach as compared to the use of heteroskedasticity-and-autocorrelation-consistent estimates is that joint hypotheses about $\boldsymbol{\delta}_N$ and $\boldsymbol{\theta}_N$ may be formulated and tested.

The TSLS estimator of model (1a) is defined as

$$\begin{aligned}
&\tilde{\tilde{\boldsymbol{\delta}}}_N = (\hat{\mathbf{Z}}_N'\mathbf{Z}_N)^{-1}\hat{\mathbf{Z}}_N'\mathbf{y}_N \text{, where} \qquad (47)\\
&\hat{\mathbf{Z}}_N = \mathbf{P}_{\mathbf{H}_N}\mathbf{Z}_N = (\mathbf{X}_N, \hat{\overline{\mathbf{Y}}}_N) \text{, and}\\
&\hat{\overline{\mathbf{Y}}}_N = \mathbf{P}_{\mathbf{H}_N}\overline{\mathbf{Y}}_N .
\end{aligned}$$

In the second step, the parameters $\rho_{s,N}$, $s=1,...,S$, σ_v^2, and σ_1^2, are estimated using the GM estimator defined by (18), based on the first step residuals $\tilde{\tilde{\mathbf{u}}}_N = \mathbf{y}_n - \mathbf{Z}_N\tilde{\tilde{\boldsymbol{\delta}}}_N$. As above these estimators are denoted as $\tilde{\rho}_{s,N}$, $s=1,...,S$, $\tilde{\sigma}_{v,N}^2$, and $\tilde{\sigma}_{1,N}^2$.

The following lemma shows that the various assumptions maintained in Section III are automatically satisfied by the TSLS estimator $\tilde{\tilde{\boldsymbol{\delta}}}_N$ and the corresponding residuals $\tilde{\tilde{\mathbf{u}}}_N$.

Lemma 1.[12]

---

[11] Kelejian, Prucha, and Yuzefovich (2004) consider the results using alternative sets of instruments in the estimation of a cross-section SARAR(1,1) model. Their Monte Carlo simulation results suggest that choosing $Q=2$ will be sufficient in many applications.

[12] Compare Kelejian and Prucha (2008) for analogous results in case of a cross-section SARAR(1,1) model and Badinger and Egger (2008b) in case of a cross-section SARAR(*R*,*S*) model.

Suppose that Assumptions 1-3 and 8-10 hold, and that $\sup_N \|\boldsymbol{\beta}_N\| \leq b < \infty$. Let $\mathbf{D}_N = -\mathbf{Z}_N$, then, the fourth moments of the elements of $\mathbf{D}_N$ are bounded uniformly in absolute value, Assumption 6 holds, and

(a) $(NT)^{1/2}(\tilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N) = (NT)^{-1/2}\mathbf{T}_N'\boldsymbol{\xi}_N + o_p(1) = (NT)^{-1/2}\mathbf{T}_{v,N}'\mathbf{v}_N + (NT)^{-1/2}\mathbf{T}_{\mu,N}'\boldsymbol{\mu}_N + o_p(1)$, where

$\boldsymbol{\xi}_N = (\mathbf{v}_N', \boldsymbol{\mu}_N')'$, $\mathbf{T}_N = (\mathbf{T}_{v,N}', \mathbf{T}_{\mu,N}')'$,

$\mathbf{T}_{v,N} = \mathbf{F}_{v,N}\mathbf{P}_N$, $\mathbf{T}_{\mu,N} = \mathbf{F}_{\mu,N}\mathbf{P}_N$,

$\mathbf{P}_N = \mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ}}(\mathbf{Q}_{\mathbf{HZ}}'\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ}})^{-1}$,

$\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')^{-1}]\mathbf{H}_N$, and

$\mathbf{F}_{\mu,N} = (\mathbf{e}_T' \otimes \mathbf{I}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')^{-1}]\mathbf{H}_N$.

(b) $(NT)^{-1/2}\mathbf{T}_N'\boldsymbol{\xi}_N = O_p(1)$;

(c) $\mathbf{P}_N = O_p(1)$ and $\tilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$ for

$\tilde{\mathbf{P}}_N = [(NT)^{-1}\mathbf{H}_N'\mathbf{H}_N]^{-1}[(NT)^{-1}\mathbf{H}_N'\mathbf{Z}_N]\{[(NT)^{-1}\mathbf{Z}_N'\mathbf{H}_N][(NT)^{-1}\mathbf{H}_N'\mathbf{H}_N]^{-1}[(NT)^{-1}\mathbf{H}_N'\mathbf{Z}_N]\}^{-1}$.

The condition $\sup_N \|\boldsymbol{\beta}_N\| \leq b < \infty$ is trivially satisfied if $\boldsymbol{\beta}_N = \boldsymbol{\beta}$. Note that (a) and (b) together imply that $\tilde{\boldsymbol{\delta}}_N$ is a $N^{1/2}$-consistent estimator of $\boldsymbol{\delta}_N$.

Regarding Assumption 4, we now have $\tilde{\mathbf{u}}_N - \mathbf{u}_N = \mathbf{D}_N\boldsymbol{\Delta}_N$, where $\mathbf{D}_N = -\mathbf{Z}_N$ and $\boldsymbol{\Delta}_N = \tilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N$. Lemma 1 shows that under Assumptions 1-3 and 8-10 the TSLS residuals automatically satisfy the conditions postulated in Assumptions 4, 6, and 7 with respect to $\mathbf{D}_N$, $\boldsymbol{\Delta}_N$, and $\mathbf{T}_N$. Hence, Theorems 1 and 2 apply to the GM estimator $\tilde{\boldsymbol{\theta}}_N$, which is based on the TSLS residuals. The lemma also establishes that the elements of $\mathbf{D}_N$ are bounded uniformly in absolute value, gives explicit expressions for $\mathbf{P}_N$ and $\tilde{\mathbf{P}}_N$, and verifies that the conditions concerning these matrices made in Theorems 3 and 4 are fulfilled. Hence, Theorems 3 and 4 cover the GM estimator $\tilde{\boldsymbol{\theta}}_N$ and the TSLS estimator $\tilde{\boldsymbol{\delta}}_N$. In particular, Theorem 4 gives the joint limiting distribution of $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ and $N^{1/2}(\tilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N)$, where $\mathbf{D}_N = -\mathbf{Z}_N$, the matrices $\mathbf{P}_N, \tilde{\mathbf{P}}_N$, $\mathbf{F}_{v,N}$, $\mathbf{F}_{\mu,N}$ are as in Lemma 1, $\tilde{\mathbf{F}}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}_{m,N}')^{+}]\mathbf{H}_N$ and

$\tilde{\mathbf{F}}_{\mu,N} = (\mathbf{e}_T' \otimes \mathbf{I}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}_{m,N}')^{+}]\mathbf{H}_N$.

We now turn to the third step of the estimation. Consider the transformed model (1b), with

$$\mathbf{y}_N^{**} = \mathbf{Z}_N^{**}\boldsymbol{\delta}_N + \mathbf{u}_N^{**}, \tag{48}$$

where

$$\mathbf{y}_N^{**} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{y}_N$$

$$\mathbf{Z}_N^{**} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{Z}_N,$$

$$\mathbf{u}_N^{**} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{u}_N = \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}\boldsymbol{\varepsilon}_N,$$

$$\mathbf{H}_N^{**} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N = \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}\mathbf{H}_N^{*},$$

$$\mathbf{H}_N^{*} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N,$$

$$\hat{\mathbf{Z}}_N^{**} = \mathbf{P}_{\mathbf{H}_N^{**}}\mathbf{Z}_N^{**} = \mathbf{H}_N^{**}(\mathbf{H}_N^{**\prime}\mathbf{H}_N^{**})^{-1}\mathbf{H}_N^{**\prime}\mathbf{Z}_N^{**}.$$

The generalized two-stages least squares (GTSLS) estimator, denoted as $\hat{\boldsymbol{\delta}}_N$, is then obtained as a two-stages least squares estimator applied to the transformed model (56), using the transformed instruments $\mathbf{H}_N^{**} = (\mathbf{I} - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})\mathbf{H}_N$, i.e.,

$$\hat{\boldsymbol{\delta}}_N = (\hat{\mathbf{Z}}_N^{**\prime}\mathbf{Z}_N^{**})^{-1}\hat{\mathbf{Z}}_N^{**\prime}\mathbf{y}_N^{**}. \tag{49}$$

The feasible generalized two-stages least squares (FGTSLS) estimator, denoted as $\hat{\tilde{\boldsymbol{\delta}}}_N$, is defined analogously, after replacing $\boldsymbol{\rho}_N$ by $\tilde{\boldsymbol{\rho}}_N$ ($\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}$ by $\tilde{\boldsymbol{\Omega}}_{\boldsymbol{\varepsilon},N}$), i.e.,

$$\hat{\tilde{\boldsymbol{\delta}}}_N = (\hat{\tilde{\mathbf{Z}}}_N^{**\prime}\tilde{\mathbf{Z}}_N^{**})^{-1}\hat{\tilde{\mathbf{Z}}}_N^{**\prime}\tilde{\mathbf{y}}_N^{**}, \tag{50}$$

where

$$\hat{\tilde{\mathbf{Z}}}_N^{**} = \mathbf{P}_{\tilde{\mathbf{H}}_N^{**}}\tilde{\mathbf{Z}}_N^{**}, \text{ with } \mathbf{P}_{\mathbf{H}_N} = \tilde{\mathbf{H}}_N^{**}(\tilde{\mathbf{H}}_N^{**\prime}\tilde{\mathbf{H}}_N^{**})^{-1}\tilde{\mathbf{H}}_N^{**\prime},$$

$$\tilde{\mathbf{H}}_N^{**} = \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N,$$

$$\tilde{\mathbf{Z}}_N^{**} = \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{Z}_N,$$

$$\tilde{\mathbf{y}}_N^{**} = \tilde{\boldsymbol{\Omega}}_{\boldsymbol{\varepsilon},N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \tilde{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{y}_N .$$

Kelejian and Prucha (2008) and Arraiz, Drukker, Kelejian and Prucha (2007) use the untransformed instrument matrix $\mathbf{H}_N$ in the FGTSLS estimation of cross-section SARAR(1,1) models. While this choice does not affect consistency, it has implications for the efficiency of the estimates. In light of (45), the ideal instruments matrix for $\overline{\mathbf{Y}}_N^{**}$ in the transformed model is given by $\mathbf{H}_N^{**}$.

The following lemma shows that the various assumptions maintained in Section III are automatically satisfied by the (feasible) generalized TSLS estimator $\tilde{\tilde{\boldsymbol{\delta}}}_N$ and the corresponding residuals. The proof is given in Appendix B.

Lemma 2.[13]

Suppose the Assumptions of Lemma 1 hold[14], and let $\hat{\tilde{\boldsymbol{\delta}}}_N$ be defined as in equation (50), where $\breve{\boldsymbol{\theta}}_N$ is any $N^{1/2}$-consistent estimator of $\boldsymbol{\theta}_N$ (such as the GM estimator $\tilde{\boldsymbol{\theta}}_N$ based on the TSLS residuals). Then

(a) $(NT)^{1/2}\boldsymbol{\Delta}_N^{**} = (NT)^{-1/2}\mathbf{T}_N^{**\prime}\boldsymbol{\xi}_N + o_p(1) = (NT)^{-1/2}\mathbf{T}_{v,N}^{**\prime}\mathbf{v}_N + (NT)^{-1/2}\mathbf{T}_{\mu,N}^{**\prime}\boldsymbol{\mu}_N + o_p(1)$, where

$\boldsymbol{\xi}_N = (\mathbf{v}_N', \boldsymbol{\mu}_N')'$, $\mathbf{T}_N^{**} = (\mathbf{T}_{v,N}^{**\prime}, \mathbf{T}_{\mu,N}^{**\prime})'$,

$\mathbf{T}_{v,N}^{**} = \mathbf{F}_{v,N}^{**}\mathbf{P}_N^{**}$, $\mathbf{T}_{\mu,N}^{**} = \mathbf{F}_{\mu,N}^{**}\mathbf{P}_N^{**}$,

$\mathbf{P}_N^{**} = \mathbf{Q}_{\mathbf{H}^{**}\mathbf{H}^{**}}^{-1}\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}}(\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}}'\mathbf{Q}_{\mathbf{H}^{**}\mathbf{H}^{**}}^{-1}\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}})^{-1}$,

$\mathbf{F}_{v,N}^{**} = (\sigma_v^{-2}\mathbf{Q}_{0,N} + \sigma_1^{-2}\mathbf{Q}_{1,N})\mathbf{H}_N^{*} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$,

$\mathbf{F}_{\mu,N}^{**} = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)]\mathbf{H}_N^{*} = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$.

(b) $(NT)^{-1/2}\mathbf{T}_N^{**\prime}\boldsymbol{\xi}_N = O_p(1)$;

---

[13] Compare Kelejian and Prucha (2008) for analogous results in case of a cross-section SARAR(1,1) model and Badinger and Egger (2008b) in case of a cross-section SARAR(*R*,*S*) model.

[14] In light of the properties maintained with respect to the matrices $(\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})$ and $\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}$, this implies that Assumptions 9 and 10 will also be satisfied for the transformed instruments $\mathbf{H}_N^{**}$.

(c) $\mathbf{P}_N^{**} = O_p(1)$ and $\breve{\mathbf{P}}_N^{**} - \mathbf{P}_N^{**} = o_p(1)$ for

$\breve{\mathbf{P}}_N^{**} = [(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}]^{-1}[(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**}]\times\{[(NT)^{-1}\breve{\mathbf{Z}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}][(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}]^{-1}[(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**}]\}^{-1}$.

In light of Lemmata 1 and 2 the joint limiting distribution of the (feasible) generalized TSLS estimator $\hat{\breve{\boldsymbol{\delta}}}_N$ and the GM estimator $\breve{\boldsymbol{\theta}}_N$ follows from Theorem 4 and the discussion thereafter, with $\boldsymbol{\Delta}_N^{**} = \hat{\breve{\boldsymbol{\delta}}}_N - \boldsymbol{\delta}_N$. The asymptotic variance-covariance matrix and its corresponding estimator are provided in Theorem 4 with the modifications as described in Remark 2 thereafter.

Note that in light of Lemma 2 the residuals $\hat{\breve{\mathbf{u}}}_N^{**} = \mathbf{y}_N - \mathbf{Z}_N\hat{\breve{\boldsymbol{\delta}}}_N = \mathbf{u}_N + \mathbf{D}_N\boldsymbol{\Delta}_N^{**}$ can be used to estimate $\boldsymbol{\theta}_N$ by the GM estimator defined by (18), where the discussion surrounding Lemma 2 would also apply here. Taking this argument one step further, $\boldsymbol{\theta}_N$ and $\boldsymbol{\delta}_N$ can also be estimated by an iterative procedure.

As a final point, note that the above theory carries over to cases where the regressor matrix $\mathbf{X}_N$ includes endogenous variables, provided that suitable instruments are available. To be more specific, let $\mathbf{X}_N = (\overline{\mathbf{X}}_N, \mathbf{E}_N)$ and $\mathbf{D}_N = -\mathbf{Z}_N = -(\overline{\mathbf{X}}_N, \mathbf{E}_N, \mathbf{Y}_N)$, where $\overline{\mathbf{X}}_N$ satisfies Assumptions 8-10 with $\mathbf{X}_N$ replaced by $\overline{\mathbf{X}}_N$ (including in the formulation of the instruments), and where $\mathbf{E}_N$ is a matrix of endogenous variables. Then, given the fourth moments of $\mathbf{D}_N$ are bounded uniformly, and Assumption 6 holds, parts (a), (b), and (c) of Lemma 1 and 2 still hold, but with $\hat{\mathbf{Z}}_N = (\overline{\mathbf{X}}_N, \mathbf{P}_{\mathbf{H}_N}\mathbf{E}_N, \mathbf{P}_{\mathbf{H}_N}\overline{\mathbf{Y}}_N)$, $\hat{\mathbf{Z}}_N^{**} = \mathbf{P}_{\mathbf{H}_N^{**}}\mathbf{Z}_N^{**}$, and $\hat{\breve{\mathbf{Z}}}_N^{**} = \mathbf{P}_{\breve{\mathbf{H}}_N^{**}}\breve{\mathbf{Z}}_N^{**}$ respectively.

V. Monte Carlo Evidence

In this section, we consider a Monte Carlos experiment for a SARAR(3,3) specification and restricted versions thereof. We assume that $\mathbf{W}_N = \mathbf{M}_N$ and that the matrix $\mathbf{X}_N$ includes two explanatory variables. Hence we have[15]

$$\mathbf{y} = \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \sum_{r=1}^{3}\lambda_r(\mathbf{I}_T \otimes \mathbf{W}_r)\mathbf{y} + \mathbf{u}, \tag{51a}$$

$$\mathbf{u} = \sum_{s=1}^{3}\rho_s(\mathbf{I}_T \otimes \mathbf{W}_s)\mathbf{u} + \boldsymbol{\varepsilon}. \tag{51b}$$

[15] For simplicity of notation, the subscript N is suppressed in the following.

We consider two sample sizes: $N = 100$ and $N = 500$. The explanatory variables $\mathbf{x}_1$ and $\mathbf{x}_2$ are generated as random draws from a standard normal distribution, scaled with a factor of five, and treated as fixed in repeated samples. Their parameters β_1 and β_2 are assumed to be unity in all Monte Carlo experiments considered.

For our basic setup of the weights matrix, we follow Kelejian and Prucha (1999) and use a binary 'up to 9 ahead and up to 9 behind' contiguity specification. This means that the elements of the time-invariant, raw weights matrix $\mathbf{W}^0$ are defined such that the *i*-th cross-section element is related to the 9 elements after it and the 9 elements before it.

The unnormalized $N \times N$ matrix $\mathbf{W}^0$ is then split up into three $N \times N$ matrices $\mathbf{W}_1^0, \mathbf{W}_2^0$, and $\mathbf{W}_3^0$, where $\mathbf{W}_1^0 + \mathbf{W}_2^0 + \mathbf{W}_3^0 = \mathbf{W}^0$. The matrices $\mathbf{W}_1^0$, $\mathbf{W}_2^0$, and $\mathbf{W}_3^0$ are specified such that they contain the elements of $\mathbf{W}_0$ for a different band of neighbours each. Otherwise, they have zero elements. We choose a design, where $\mathbf{W}_1^0$ corresponds to an 'up to 3 ahead and up to 3 behind' specification, $\mathbf{W}_2^0$ corresponds to a '4 to 6 ahead and 4 to 6 behind' specification, and $\mathbf{W}_3^0$ corresponds to a '7 to 9 ahead and 7 to 9 behind' specification. $\mathbf{W}_1^0$, $\mathbf{W}_2^0$, and $\mathbf{W}_3^0$ have typical elements $w_{1,ij}^0$, $w_{2,ij}^0$, and $w_{3,ij}^0$, respectively, where subscripts i and j indicate that the corresponding element captures the possible contiguity of unit i with j. $w_{1,ij}^0$, $w_{2,ij}^0$, and $w_{3,ij}^0$ are either unity or zero. By design, at most one of the three elements, $w_{1,ij}^0$, $w_{2,ij}^0$, or $w_{3,ij}^0$, can be unity. The final weights matrices $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{W}_3$ are obtained by separately row-normalizing $\mathbf{W}_1^0$, $\mathbf{W}_2^0$, and $\mathbf{W}_3^0$, that is, by dividing their typical elements $w_{1,ij}^0$, $w_{2,ij}^0$, and $w_{3,ij}^0$ through the corresponding row sum, respectively.

With three row-normalized matrices $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{W}_3$, the parameter space for $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$ must satisfy $0 \le |\lambda_1| + |\lambda_2| + |\lambda_3| < 1$ and $0 \le |\rho_1| + |\rho_2| + |\rho_3| < 1$. We consider 3 parameter constellations. In constellation (1) there is third order spatial dependence in both the dependent variable and the disturbances, which is non-increasing in the order of neighbourhood, i.e., $\lambda_1 \ge \lambda_2 \ge \lambda_3$ and $\rho_1 \ge \rho_2 \ge \rho_3$. In parameter constellation (2), there is first order spatial dependence in both $\mathbf{y}$ and $\mathbf{u}$. Finally, parameter constellation (3) considers zero dependence parameters for all spatial lags in $\mathbf{y}$ and $\mathbf{u}$, i.e., a non-spatial model.

< Table 1 here >

Regarding the choice of instruments, we include linearly independent terms of up to the second order in equation (30b). In particular, the matrix of untransformed instruments $\mathbf{H}$ contains 18 columns and is given by

$$\mathbf{H} = (\mathbf{X}, \mathbf{W}_1\mathbf{X}, \mathbf{W}_2\mathbf{X}, \mathbf{W}_3\mathbf{X}, \mathbf{W}_1^2\mathbf{X}, \mathbf{W}_2^2\mathbf{X}, \mathbf{W}_3^2\mathbf{X}, \mathbf{W}_{12}\mathbf{X}, \mathbf{W}_{23}\mathbf{X}), \tag{52}$$

where $\mathbf{W}_{ij} = \mathbf{W}_i\mathbf{W}_j$.

We assume further that the error components v_{it} and μ_{it} are drawn from a standard normal distribution with zero mean and unit variance, i.e., $v_{it} = \zeta_{v,it}$ and $\mu_{it} = \zeta_{\mu,i}$ where each $\zeta_{v,it}$ and $\zeta_{\mu,i}$ are i.i.d. $N(0,1)$.

For each Monte Carlo experiment, we consider 1000 draws. To ensure comparability, the same draws of $\zeta_{v,it}$ and $\zeta_{\mu,it}$ are used for each parameter constellation. Results for the estimates of $\rho_{1,N}, \rho_{2,N}$, and $\rho_{3,N}$ are obtained by the GM estimator defined in equation (18), using the optimal weighting matrix under normality as given in equation $(\tilde{\boldsymbol{\Psi}}_N^{\circ})^{-1}$. The estimates reported for the regression parameters are FGTSLS estimates as defined in (50) using the transformed set of instruments $\tilde{\mathbf{H}}^{**}$.

For each single coefficient, we report the average bias and root mean squared error for each parameter constellation and the rejection rates for the test that the coefficient is equal to the true parameter value. Under parameter constellation (2) we also test the SARAR(3,3) against the SARAR(1,1) model, using $H_0^{\lambda,\rho,*}: \lambda_2 = \lambda_3 = \rho_2 = \rho_3 = 0$. For the non-spatial model under parameter constellation (3), we report results for the tests of the joint hypothesis $H_0^{\lambda,\rho}: \lambda_1 = \lambda_2 = \lambda_3 = \rho_1 = \rho_2 = \rho_3 = 0$.

Using Theorem 4, the approximation of the small sample distribution of $\tilde{\mathbf{q}}$ is given by $\tilde{\mathbf{q}} \sim N(\mathbf{q}, \boldsymbol{\mathfrak{Q}})$, where $\mathbf{q} = (\lambda_1, \lambda_2, \lambda_3, \beta_1, \beta_2, \rho_1, \rho_2, \rho_3)'$ and $\boldsymbol{\mathfrak{Q}} = Var(\tilde{\mathbf{q}})$, which can be estimated using $\tilde{\boldsymbol{\mathfrak{Q}}} = N^{-1}\tilde{\boldsymbol{\Omega}}_o$. Tests referring to a single parameter are carried out using a standard t-test, e.g., $H_0^{\rho}: \tilde{\rho}_1 = \rho_1$ is tested using $t = \dfrac{\tilde{\rho}_1 - \rho_1}{\tilde{\sigma}_{\tilde{\rho}_1}}$, where $\sigma_{\tilde{\rho}_1}$ is the corresponding main diagonal element of $\tilde{\boldsymbol{\mathfrak{Q}}}$.

Tests regarding joint hypotheses are carried out using Wald tests. Generally, we test $H_0: \mathfrak{R}\mathbf{q} - \mathfrak{t} = \mathbf{0}$ against $H_1: \mathfrak{R}\mathbf{q} - \mathfrak{t} \neq \mathbf{0}$. Define the discrepancy vector: $\mathfrak{m} = \mathfrak{R}\tilde{\mathbf{q}} - \mathfrak{t}$. The null hypothesis can the be tested using $\mathfrak{m}'(\mathfrak{R}\tilde{\boldsymbol{\mathfrak{Q}}}\mathfrak{R}')^{-1}\mathfrak{m} \sim \chi^2_{\mathfrak{G}}$, where $\mathfrak{G}$ is the number of restrictions, i.e., the number of rows of $\mathfrak{R}$ (e.g., Greene, 2003, pp. 95, 487). In the present context, we have

for $H_0^{\lambda,\rho,*}: \lambda_2 = \lambda_3 = \rho_2 = \rho_3 = 0$, $\mathfrak{R} = \begin{bmatrix} 0 & \mathbf{I}_2 & 0 & 0 & 0 & \mathbf{0} \\ 0 & \mathbf{0} & 0 & 0 & 0 & \mathbf{I}_2 \end{bmatrix}$ and $\mathfrak{G} = 4$;

for $H_0^{\lambda,\rho}: \lambda_1 = \lambda_2 = \lambda_3 = \rho_1 = \rho_2 = \rho_3 = 0$, $\mathfrak{R} = \begin{bmatrix} \mathbf{I}_3 & 0 & 0 & \mathbf{0} \\ \mathbf{0} & 0 & 0 & \mathbf{I}_3 \end{bmatrix}$ and $\mathfrak{G} = 6$.

Table 2 reports the results of the Monte Carlo analysis for the two sample sizes considered.[16] In terms of bias and RMSE, the estimator performs well, even in the small sample with $N = 100$. On average over all parameter constellation the bias and RMSE amount to 0.0007 and 0.0229 for the estimates of $\boldsymbol{\lambda} = (\lambda_1,...,\lambda_3)'$ and to 0.0054 and 0.1096 for the estimates of $\boldsymbol{\rho} = (\rho_1,...,\rho_3)'$. With an average rejection rate of 0.0082, the performance of the single hypotheses tests referring to $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$ is satisfactory. The actual size of the joint hypotheses tests, however, differs significantly from the nominal size with an average rejection rate of 0.1395.

< Table 2 >

However, performance improves quickly with growing sample size. For $N = 500$, the bias virtually disappears and the average RMSE of the estimates of $\boldsymbol{\lambda} = (\lambda_1,...,\lambda_3)'$ shrinks to 0.0010, that of the estimates of $\boldsymbol{\rho} = (\rho_1,...,\rho_3)'$ shrinks to 0.0440. Also, the size of the tests improves and approaches the nominal size of 5 percent. Regarding the GM estimates of $\boldsymbol{\rho}$, the average size of the tests involving only one parameter amounts to 0.0089, that for the FGTSLS estimates of $\boldsymbol{\lambda}$ to 0.053. The average size of the joint hypothesis amounts to 0.084 for the joint tests.

Overall, the Monte Carlo experiments illustrate that the proposed estimators work reasonably well in terms of bias and RMSE, even in very small samples. Regarding the estimates of the variance-covariance matrix of the parameter estimates and implied tests of single and joint hypotheses, some care is warranted in the interpretation of the results in small samples, though the difference to the true size of the tests is moderate at least for the single hypothesis tests. This suggest that in small samples it might be worth exploiting additional moment conditions as outlined in footnote 3. As the sample size increases, the rejection rates of the single and joint tests converge reasonably quickly to the true size such that they may be recommended for specification tests about the lag- and error-structure and the order of spatial dependence in medium to large samples.

[16] Results for the variances of the error components are very similar to those of the GM estimates of the spatial regressive parameters of the disturbance process. The detailed results are thus omitted for the sake of brevity. The only notable difference is that the rejection rates for the estimates of σ_1^2 are worse with an average value of 0.175 for $N = 100$ and 0.138 for $N = 500$.

VI. Conclusions and Suggestions for Future Research

This paper derives GM and FGTSLS estimators for the parameters of a SARAR(*R*,*S*) model allowing the applied econometrician to study the strength and pattern of spatial interdependence more flexibly than existing SARAR(1,1) models. We also provide a detailed study of the asymptotic properties of the proposed two-step GM-FGLS estimators of the model parameters and derive their joint asymptotic distribution. This enables tests of the fairly general SARAR(*R*,*S*) model against restricted alternatives such as SARAR(0,*S*) and SARAR(*R*,0) or SARAR(1,1) with panel data.

One suggestion for future research is to extend the analysis of tests towards a study of conditional and unconditional tests on the relevance of error components and spatial interaction. In particular, a comprehensive Monte Carlo study of GM estimators using alternative weighting schemes of the moments and alternative distributional assumptions should be instructive for related applied research.

References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Boston: Kluwer, Academic Publishers.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2), 153-166.

Arbia, G., Basile, R., and Piras, G. (2005), Using spatial panel data in modeling regional growth and convergence. ISAE Working Paper no. 55, Rome.

Arraiz, I., Drukker, D.M., Kelejian, H., and Prucha, I. (2007). A spatial Cliff-Ord-type model with heteroskedastic innovations: Small and large sample results. Unpublished manuscript.

Audretsch, D.B. and Feldmann, M.P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86, 630-640.

Badinger, H. and Egger, P. (2008a). Horizontal versus vertical interdependence in multinational activity. CESifo Working Paper no. 2327, Munich.

Badinger, H. and Egger, P. (2008b). GM estimation of higher-order spatial autoregressive processes in cross-section models with heteroskedastic disturbances. CESifo Working Paper no. 2356, Munich.

Baltagi, B.H., Egger, P., and Pfaffermayr, M. (2007). Estimating models of complex FDI: Are there third-country effects? *Journal of Econometrics*, 140(1), 260-281.

Baltagi, B.H., Egger, P., and Pfaffermayr, M. (2008). A generalized spatial panel data model with random effects. Working paper, University of Syracuse.

Baltagi, B.H. and Li, D. (2001). LM test for functional form and spatial error correlation. *International Regional Science Review*, 24, 194-225.

Baltagi, B.H., and Li, D. (2004). Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer, Berlin, pp. 283–295.

Baltagi, B.H., Song, S.H., and Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117, 123-150.

Baltagi, B.H., Song, S.H., Jung, B.C., and Koh, W. (2007). Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics*, 140, 5-51.

Bell, K.P. and Bockstael, N.E. (2000). Applying the generalized-moments estimation apporoach to spatial problems involving microlevel data. *The Review of Economics and Statistics,* 82(1), 72–82.

Besley, T. and Case, A. (1995). Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition. *American Economic Review*, 85, 25-45.

Case, A., Hines Jr., J. and Rosen, H. (1993). Budget spillovers and fiscal policy independence: Evidence from the States. *Journal of Public Economics*, 52, 285-307.

Cliff, A. and Ord, J. (1973). *Spatial Autocorrelation*. London: Pion, 1973.

Cliff, A. and Ord, J. (1981). *Spatial Processes, Models and Applications*. London: Pion, 1981.

Cohen, J.P. and Morrison Paul, C. (2007). The impacts of transportation infrastructure on property values: A higher-order spatial econometrics approach. *Journal of Regional Science*, 47(3), 457-478.

Cohen, J.P. and Morrison Paul, C.J. (2004). Public infrastructure investment, interstate spatial spillovers, and manufacturing costs. *The Review of Economics and Statistics,* 86(2), 551-560.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92, 1-45.

Druska, V. and Horrace, W.C. (2004), Generalized moments estimation for spatial panel data: Indonesian rice farming. *American Journal of Agricultural Economics*, 86, 185-198.

Egger, P., Pfaffermayr, M. and Winner, H. (2005). An unbalanced spatial panel data approach to US state tax competition. *Economics Letters*, 88, 329-335.

Gilbert, S. (2002). Testing the distribution of error components in panel data models. *Economics Letters*, 77, 47-53.

Greene, W.H. (2003). *Econometric Analysis*, fifth edition. Pearson, Upper Saddle River, New Jersey.

Holtz-Eakin, D. (1994). Public sector capital and the productivity puzzle. *Review of Economics and Statistics*, 76, 12-21.

Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.

Kapoor, M., Kelejian, H.H., and Prucha, I.R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140, 97-130.

Kelejian, H. and Robinson, D. (1992). Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures. *Regional Science and Urban Economics*, 22, 317–331.

Kelejian, H.H. and Prucha, I.R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17, 99-121.

Kelejian, H.H. and Prucha, I.R. (1999). A generalized moments estimator for the autoregresssive parameter in a spatial model. *International Economic Review*, 40, 509-533.

Kelejian, H.H. and Prucha, I.R. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics*, 118, 27-50.

Kelejian, H.H. and Prucha, I.R. (2007). HAC Estimation in a Spatial Framework. *Journal of Econometrics*, 140(1), 131-154.

Kelejian, H.H. and Prucha, I.R. (2008). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, forthcoming.

Kelejian, H.H., Prucha, I.R. and Yuzefovich, E. (2004). Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small sample results. In: LeSage, J. and Pace, K. (eds.), *Advances in Econometrics: Spatial and Spatiotemporal Econometrics*. Elsevier, New York, 163-198.

Korniotis, G. (2008). Estimating panel models with internal and external habit formation. *Journal of Business and Economic Statistics*, forthcoming.

Lee, L.F. and Liu, X. (2008). Efficient GMM estimation of high order spatial autoregressive models. Working Paper, Department of Economics, Ohio State University.

Lee, L.F. and Yu, J. (2008). Estimation of spatial autoregressive panel data models with fixed effects. Working Paper, Department of Economics, Ohio State University.

Mittelhammer, R.C. (1996). *Mathematical Statistics for Economics and Business*. New York: Springer.

Pinkse, J. and Slade, M.E. (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85, 125-154.

Pinkse, J., Slade, M.E., and Brett, C. (2002). Spatial price competition: A semiparametric approach. *Econometrica*, 70, 1111-1153.

Pötscher, B.M. and Prucha, I.R. (1997). *Dynamic Nonlinear Econometric Models*, *Asymptotic Theory*. New York: Springer.

Rao, C.R. (1973). *Linear Statistical Inference and its Applications*, 2nd edition. New York: Wiley.

Resnik, S. (1999). *A Probability Path*. Boston: Birkhäuser.

Shroder, M. (1995). Games the States don't play: Welfare benefits and the theory of fiscal federalism. *Review of Economics and Statistics*, 77, 183-191.

Topa, G. (2001). Social interactions, local spillovers and unemployment. *Review of Economic Studies*, 68, 261-295.

APPENDIX

In the following, we sketch the proofs of Theorems 1-4. They build on analogous proofs by Kelejian and Prucha (2008) for a cross-sectional SARAR(1,1) and Badinger and Egger (2008b) for a cross-section SARAR(*R*,*S*) model, as well as analogous proofs for a panel SARAR(0,1) model with nonstochastic regressors by Kapoor, Kelejian, and Prucha (2007). The full details of the proofs are given in a technical appendix to this paper.

APPENDIX A

Notation

We adopt the standard convention to refer to matrices and vectors with acronyms in boldface. Let $\mathbf{A}_N$ denote some matrix. Its elements are referred to as $a_{ij,N}$; $\mathbf{a}_{i.,N}$ and $\mathbf{a}_{.i,N}$ denote the *i*-th row and the *i*-th column of $\mathbf{A}_N$ respectively. If $\mathbf{A}_N$ is a square matrix, $\mathbf{A}_N^{-1}$ denotes its inverse; if $\mathbf{A}_N$ is singular, $\mathbf{A}_N^{+}$ denotes its generalized inverse. If $\mathbf{A}_N$ is a square, symmetric and positive definite matrix, $\mathbf{A}_N^{1/2}$ denotes the unique positive definite square root of $\mathbf{A}_N$ and $\mathbf{A}_N^{-1/2}$ denotes $(\mathbf{A}_N^{-1})^{1/2}$. The (submultiplicative) matrix norm $\|\cdot\|$ is defined as $\|\mathbf{A}_N\| = [Tr(\mathbf{A}_N'\mathbf{A}_N)]^{1/2}$. Finally, unless stated otherwise, for expressions involving sums over elements of vectors or matrices that are stacked over all time periods, we adopt the convention to use single indexation *i*, running from $i = 1,...,NT$, to denote elements of the stacked vectors or matrices.[17]

Remark A.1

i) Definition of row and column sum boundedness (compare Kapoor, Kelejian, and Prucha, 2007, p. 99): Let $\mathbf{B}_N, N \geq 1$, be some sequence of $NT \times NT$ matrices with T some fixed positive integer. We will then say that the row and column sums of the (sequence of) matrices $\mathbf{B}_N$ are bounded uniformly in absolute value, if there exists a constant $c < \infty$, which does not depend on *N*, such that

$$\max_{1 \leq i \leq NT} \sum_{j=1}^{NT} |b_{ij,N}| \leq c \text{ and } \max_{1 \leq j \leq NT} \sum_{i=1}^{NT} |b_{ij,N}| \leq c \text{ for all } N \geq 1.$$

The following results will be repeatedly used in the subsequent proofs.

---

[17] Take the vector $\mathbf{u}_N = [\mathbf{u}_N'(1),...,\mathbf{u}_N'(T)]$, for example. Using indexation $i = 1,...,NT$, the elements $u_{i,N}, i = 1,...,N$ refer to $t = 1$, elements $u_{i,N}, i = N+1,...,2N$ refer to $t = 2$, etc., and elements $u_{i,N}, i = (T-1)N+1,...,NT$ refer to $t = T$. The major advantage of this notation is that it avoids the use of double indexation for the cross-section and time dimension. Moreover, it allows us the invoke several results referring to the case of a single cross-section, which still apply to the case of T stacked cross-sections.

ii) Let $\mathbf{R}_N$ be a (sequence of) $N \times N$ matrices whose row and column sums are bounded uniformly in absolute value, and let $\mathbf{S}$ be some $T \times T$ matrix (with $T \geq 1$ fixed). Then the row and column sums of the matrix $\mathbf{S} \otimes \mathbf{R}_N$ are bounded uniformly in absolute value (compare Kapoor, Kelejian, and Prucha, 2007, p. 118).

iii) If $\mathbf{A}_N$ and $\mathbf{B}_N$ are (sequences of) $NT \times NT$ matrices (with $T \geq 1$ fixed), whose row and column sums are bounded uniformly in absolute value, then so are the row and column sums of $\mathbf{A}_N\mathbf{B}_N$ and $\mathbf{A}_N + \mathbf{B}_N$. If $\mathbf{Z}_N$ is a (sequence of) $NT \times P$ matrices whose elements are bounded uniformly in absolute value, then so are the elements of $\mathbf{A}_N\mathbf{Z}_N$ and $(NT)^{-1}\mathbf{Z}_N'\mathbf{A}_N\mathbf{Z}_N$. Of course, this also covers the case $(NT)^{-1}\mathbf{Z}_N'\mathbf{Z}_N$ for $\mathbf{A}_N = \mathbf{I}_{NT}$ (compare Kapoor, Kelejian, and Prucha, 2007, p. 119).

iv) Suppose that the row and columns sums of the $NT \times NT$ matrices $\mathbf{A}_N = (a_{ij,N})$ are bounded uniformly in absolute value by some finite constant c_A; then $\sum_{i=1}^{NT} \left| a_{ij,N} \right|^q \leq c_A^q$ for $q > 1$ (see Kelejian and Prucha, 2008, Remark C.1).

v) Let $\boldsymbol{\xi}_N$ and $\boldsymbol{\eta}_N$ be $NT \times 1$ random vectors (with $T \geq 1$ fixed), where, for each *N*, the elements are independently distributed with zero mean and finite variances. Then the elements of $(NT)^{-1/2}\mathbf{Z}_N'\boldsymbol{\xi}_N$ are $O_p(1)$ and $(NT)^{-1}\boldsymbol{\xi}_N'\mathbf{A}_N\boldsymbol{\eta}_N$ is $O_p(1)$ (compare Kelejian and Prucha, 2004, Remark A.1).[18]

vi) Let $\boldsymbol{\zeta}_N$ be a $NT \times 1$ random vector (with $T \geq 1$ fixed), where, for each *N*, the elements are distributed with zero mean and finite fourth moments. Let $\boldsymbol{\pi}_N$ be some nonstochastic $NT \times 1$ vector, whose elements are bounded uniformly in absolute value and let $\boldsymbol{\Pi}_N$ be a $NT \times NT$ nonstochastic matrix whose row and column sums are bounded uniformly in absolute value. Define the column vector $\mathbf{d}_N = \boldsymbol{\pi}_N + \boldsymbol{\Pi}_N\boldsymbol{\zeta}_N$. It follows that the elements of $\mathbf{d}_N$ have finite fourth moments. (Compare Kelejian and Prucha, 2008, Lemma C.2, for the case $T = 1$ and independent elements of $\boldsymbol{\zeta}_N$.)[19]

---

[18] Kelejian and Prucha (2004) consider the case $T = 1$ and where the elements of $\boldsymbol{\xi}_N$ and $\boldsymbol{\eta}_N$ are identically distributed. Obviously, the results also holds for (fixed) $T \geq 1$ and under heteroskedasticity, as long as the variances of the elements of $\boldsymbol{\xi}_N$ and $\boldsymbol{\eta}_N$ are bounded uniformly in absolute value.

[19] The extension to (fixed) $T \geq 1$ is obvious. Independence of the elements of $\boldsymbol{\zeta}_N$ is not required for the result to hold. The fourth moments of the elements of $\mathbf{d}_N = \boldsymbol{\pi}_N + \boldsymbol{\Pi}_N\boldsymbol{\zeta}_N$ are given by $E(\pi_{i,N} + \sum_{j=1}^{NT} \pi_{ij,N}\zeta_{j,N})^4 \leq 2^4 E[\pi_{i,N}^4 + (\sum_{j=1}^{NT} \pi_{ij,N}\zeta_{j,N})^4]$

Remark A2.

The matrices $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ have the following properties (see Kapoor, Kelejian, and Prucha, 2007, p. 101):

$$tr(\mathbf{Q}_{0,N}) = N(T-1),\ tr(\mathbf{Q}_{1,N}) = N,\ \mathbf{Q}_{0,N}(\mathbf{e}_T \otimes \mathbf{I}_N) = \mathbf{0},\ \mathbf{Q}_{1,N}(\mathbf{e}_T \otimes \mathbf{I}_N) = (\mathbf{e}_T \otimes \mathbf{I}_N),$$
$$\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N = \mathbf{Q}_{0,N}\mathbf{v}_N,\ \mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N = (\mathbf{e}_T \otimes \mathbf{I}_N)\boldsymbol{\mu}_N + \mathbf{Q}_{1,N}\mathbf{v}_N,\ (\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{0,N} = \mathbf{Q}_{0,N}(\mathbf{I}_T \otimes \mathbf{D}_N),$$
$$(\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{1,N} = \mathbf{Q}_{1,N}(\mathbf{I}_T \otimes \mathbf{D}_N),\ tr[(\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{0,N}] = (T-1)tr(\mathbf{D}_N),$$
$$tr[(\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{1,N}] = tr(\mathbf{D}_N),$$

where $\mathbf{D}_N$ is an arbitrary $N \times N$ matrix. Obviously, the row and column sums of $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ are bounded uniformly in absolute value.

APPENDIX B

The following lemma will be repeatedly used in the subsequent proofs.

Lemma B.1[20]

Let $\mathbf{A}_N$ be some nonstochastic $NT \times NT$ matrix (with T fixed), whose row and column sums are bounded uniformly in absolute value. Let $\mathbf{u}_N$ be defined by (2c) and $\tilde{\mathbf{u}}_N$ be a predictor for $\mathbf{u}_N$. Suppose that Assumptions 1 to 4 hold. Then

(a) $N^{-1}E\left|\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N\right| = O(1)$, $Var(N^{-1}\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N) = o(1)$,

and $N^{-1}(\tilde{\mathbf{u}}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N) - N^{-1}E(\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N) = o_p(1)$.

(b) $N^{-1}E\left|\mathbf{d}'_{.j,N}\mathbf{A}_N\mathbf{u}_N\right| = O(1)$, $j = 1,...,P$, where $\mathbf{d}_{.j,N}$ is the j-th column of the $NT \times P$ matrix $\mathbf{D}_N$, and $N^{-1}\mathbf{D}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N - N^{-1}E(\mathbf{D}'_N\mathbf{A}_N\mathbf{u}_N) = o_p(1)$.

(c) If furthermore Assumption 6 holds, then

$N^{-1/2}\tilde{\mathbf{u}}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N = N^{-1/2}\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N + \boldsymbol{\alpha}'_N N^{1/2}\boldsymbol{\Delta}_N + o_p(1)$ with $\boldsymbol{\alpha}_N = N^{-1}E[\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\mathbf{u}_N]$.

In light of (b), we have $\boldsymbol{\alpha}_N = O(1)$ and $N^{-1}\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\tilde{\mathbf{u}}_N - \boldsymbol{\alpha}_N = o_p(1)$.

The proof of Lemma B.1 is given in the technical appendix.

$\leq 2^4[\pi_{i,N}^4 + \sum_{j=1}^{NT}\sum_{k=1}^{NT}\sum_{l=1}^{NT}\sum_{m=1}^{NT}\left|\pi_{ij,N}\right|\left\|\pi_{ik,N}\right\|\left\|\pi_{il,N}\right\|\left\|\pi_{im,N}\right|E\left|\zeta_{j,N}\right\|\zeta_{k,N}\left\|\zeta_{l,N}\right\|\zeta_{m,N}\right|] \leq K < \infty$, by Hölder's inequality as long as the fourth moments of the elements of ζ_N are bounded uniformly.

[20] Compare Lemma C.1 in Kelejian and Prucha (2008) for the case of a cross-sectional SARAR(1,1) model and Lemma C.1 in Badinger and Egger (2008b) for the case of a cross-sectional SARAR(*R*,*S*) model.

Proof of Theorem 1a. Consistency of Initial GM Estimator $\hat{\boldsymbol{\theta}}_N^0$

The objective function of the nonlinear least squares estimator in (17a) and its nonstochastic counterpart are given by

$$\mathbf{R}_N^0(\omega,\underline{\boldsymbol{\theta}}^0) = (\tilde{\boldsymbol{\gamma}}_N^0 - \tilde{\boldsymbol{\Gamma}}_N^0\mathfrak{b}_N^0)'(\tilde{\boldsymbol{\gamma}}_N^0 - \tilde{\boldsymbol{\Gamma}}_N^0\mathfrak{b}_N^0) \text{ and} \tag{B.1a}$$

$$\overline{\mathbf{R}}_N^0(\underline{\boldsymbol{\theta}}^0) = (\boldsymbol{\gamma}_N^0 - \boldsymbol{\Gamma}_N^0\underline{\mathfrak{b}}^0)'(\boldsymbol{\gamma}_N^0 - \boldsymbol{\Gamma}_N^0\underline{\mathfrak{b}}^0). \tag{B.1b}$$

Since $\boldsymbol{\gamma}_N^0 - \boldsymbol{\Gamma}_N^0\mathfrak{b}_N^0 = 0$, we have $\overline{R}_N^0(\boldsymbol{\theta}_N^0) = 0$, i.e., $\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) = 0$ at the true parameter vector $\boldsymbol{\theta}_N^0 = (\rho_1,...,\rho_S,\sigma_v^2)'$. Hence,

$$\overline{\mathbf{R}}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{\mathbf{R}}_N^0(\boldsymbol{\theta}^0) = (\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'\boldsymbol{\Gamma}_N^{0\prime}\boldsymbol{\Gamma}_N^0(\underline{\mathfrak{b}}_N^0 - \mathfrak{b}_N^0). \tag{B.2}$$

In light of Rao (1973, p. 62) and Assumption 5, it follows that:

$$\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0) \geq \lambda_{\min}(\boldsymbol{\Gamma}_N^{0\prime}\boldsymbol{\Gamma}_N^0)(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0) \text{ and}$$

$$\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0) \geq \lambda_*(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0).$$

By the properties of the norm $\|\mathbf{A}\| = [tr(\mathbf{AA})]^{1/2}$, we have $\left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\|^2 \leq (\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)$ such that $\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0) \geq \lambda_*\left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\|^2$. Hence, for every $\varepsilon > 0$

$$\lim_{N\to\infty} \inf_{\{\boldsymbol{\theta}^0 : \left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\| \geq \varepsilon\}} [\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0)] \geq \inf_{\{\boldsymbol{\theta}^0 : \left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\| \geq \varepsilon\}} \lambda_*\left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\|^2 = \lambda_*\varepsilon^2 > 0, \tag{B.3}$$

which proves that the true parameter $\boldsymbol{\theta}^0$ is identifiable unique.

Next, let $\mathbf{F}_N^0 = (\tilde{\boldsymbol{\gamma}}_N^0, -\tilde{\boldsymbol{\Gamma}}_N^0)$ and $\boldsymbol{\Phi}_N^0 = (\boldsymbol{\gamma}_N^0, -\boldsymbol{\Gamma}_N^0)$. The objective function and its nonstochastic counterpart can then be written as

$$R_N^0(\underline{\boldsymbol{\theta}}^0) = (1,\underline{\mathfrak{b}}^{0\prime})\mathbf{F}_N^{0\prime}\mathbf{F}_N^0(1,\underline{\mathfrak{b}}^{0\prime})' \text{ and}$$

$$\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) = (1,\underline{\mathfrak{b}}^{0\prime})\boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0(1,\underline{\mathfrak{b}}^{0\prime})'.$$

Hence for $\underline{\boldsymbol{\rho}} \in [-\mathbf{a}^\rho, \mathbf{a}^\rho]$[21] and $\underline{\sigma}_v^2 \in [0,b]$ it holds that

$$\left\|R_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\underline{\boldsymbol{\theta}}^0)\right\| = \left\|(1,\underline{\mathfrak{b}}^{0\prime})(\mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0)(1,\underline{\mathfrak{b}}^{0\prime})'\right\|.$$

Moreover, since the norm $\|\cdot\|$ is submultiplicative, i.e., $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$, we have

[21] This should be read as $\underline{\rho}_s \in [-a^\rho, a^\rho]$ for all $s = 1,...,S$.

$$\left\| R_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) \right\| \leq \left\| \mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0 \right\| \left\| (1, \underline{\mathbf{b}}_N^{0\prime}) \right\|^2$$

$$\leq \left\| \mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0 \right\| [1 + S(a^\rho)^2 + \frac{2S + S(S-1)}{2}(a^\rho)^4 + b^2].$$

It is readily observed from (16), that the elements of the matrices $\boldsymbol{\gamma}_N^0$ and $\boldsymbol{\Gamma}_N^0$ are all of the form $\mathbf{u}_N'\mathbf{A}_N\mathbf{u}_N$, where $\mathbf{A}_N$ are nonstochastic $NT \times NT$ matrices (with T fixed), whose row and column sums are bounded uniformly in absolute value. In light of Lemma B.1, the elements of $\boldsymbol{\Phi}_N^0$ are O(1) and it follows that $\left\| \mathbf{F}_N^0 - \boldsymbol{\Phi}_N^0 \right\| \xrightarrow{p} 0$ and $\left\| \mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0 \right\| \xrightarrow{p} 0$ as $N \to \infty$. As a consequence, we have (for finite S)

$$\sup_{\underline{\boldsymbol{\rho}} \in [-\mathbf{a}^\rho, \mathbf{a}^\rho], \underline{\sigma}_v^2 \in [0,b]} \left| R_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) \right| \leq \left\| [\mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0] \right\| [1 + S(a^\rho)^2 + \frac{S(S-1)}{2}(a^\rho)^4 + b^2] \xrightarrow{p} 0 \text{ as } N \to \infty \tag{B.4}$$

Together with identifiable uniqueness, the consistency of $\tilde{\boldsymbol{\theta}}_N^0 = (\tilde{\rho}_{1,N}, ..., \tilde{\rho}_{S,N}, \tilde{\sigma}_{v,N}^2)$ now follows directly from Lemma 3.1 in Pötscher and Prucha (1997).

Having proved that the estimators $\tilde{\rho}_{1,N} ..., \tilde{\rho}_{S,N}, \tilde{\sigma}_{v,N}^2$ are consistent for $\rho_{1,N} ..., \rho_{S,N}, \sigma_v^2$, we now show that σ_1^2 can be estimated consistently from the last line $(4S+2)$ of equation system (12), using

$$\begin{aligned} \tilde{\sigma}_{1,N}^2 = {} & \tilde{\gamma}_{4S+2,N} - \tilde{\gamma}_{4S+2,1,N}\tilde{\rho}_{1,N} - ... - \tilde{\gamma}_{4S+2,S,N}\tilde{\rho}_{S,N} - \tilde{\gamma}_{4S+2,S+1,N}\tilde{\rho}_{1,N}^2 \\ & - \tilde{\gamma}_{4S+2,2S,N}\tilde{\rho}_{S,N}^2 - \tilde{\gamma}_{4S+2,2S+1,N}\tilde{\rho}_{1,N}\tilde{\rho}_{2,N} - ... - \tilde{\gamma}_{4S+2,2S+S(S-1)/2,N}\tilde{\rho}_{S-1,N}\tilde{\rho}_{S,N}. \end{aligned} \tag{B.5}$$

Since $\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \mathbf{b}_N^0 = 0$, we have

$$\begin{aligned} & \tilde{\sigma}_1^2 - \sigma_1^2 = (\tilde{\gamma}_{4S+2,N} - \gamma_{4S+2,N}) - (\tilde{\gamma}_{4S+2,1,N} - \gamma_{4S+2,1,N})\tilde{\rho}_{1,N} - ... - (\tilde{\gamma}_{4S+2,S,N} - \gamma_{4S+2,S,N})\tilde{\rho}_{S,N} \\ & - (\tilde{\gamma}_{4S+2,S+1,N} - \gamma_{4S+2,S+1,N})\tilde{\rho}_{1,N}^2 - ... - (\tilde{\gamma}_{4S+2,2S,N} - \gamma_{4S+2,2S,N})\tilde{\rho}_{S,N}^2 \\ & - (\tilde{\gamma}_{4S+2,2S+1,N} - \gamma_{4S+2,2S+1,N})\tilde{\rho}_{1,N}\tilde{\rho}_{2,N} - ... - (\tilde{\gamma}_{4S+2,2S+S(S-1)/2,N} - \gamma_{4S+2,2S+S(S-1)/2,N})\tilde{\rho}_{S-1,N}\tilde{\rho}_{S,N} \\ & - \gamma_{4S+2,1,N}(\tilde{\rho}_{1,N} - \rho_{1,N}) - ... - \gamma_{4S+2,S,N}(\tilde{\rho}_{S,N} - \rho_{S,N}) \\ & - \gamma_{4S+2,S+1,N}(\tilde{\rho}_{1,N}^2 - \rho_{1,N}^2) ... - \gamma_{4S+2,2S,N}(\tilde{\rho}_{S,N}^2 - \rho_{S,N}^2) \\ & - \gamma_{4S+2,2S+1,N}(\tilde{\rho}_{1,N}\tilde{\rho}_{2,N}^2 - \rho_{1,N}\rho_{2,N}) ... - \gamma_{4S+2,2S+S(S-1)/2,N}(\tilde{\rho}_{S-1,N}\tilde{\rho}_{S,N} - \rho_{S-1,N}\rho_{S,N}). \end{aligned} \tag{B.6}$$

Since $\left\| \mathbf{F}_N^0 - \boldsymbol{\Phi}_N^0 \right\| \xrightarrow{p} 0$ as $N \to \infty$ and the elements of $\boldsymbol{\Phi}_N$ are O(1) it follows from the consistency of $\tilde{\rho}_{1,N}, ..., \tilde{\rho}_{S,N}$ that $\tilde{\sigma}_{1,N}^2 - \sigma_1^2 \xrightarrow{p} 0$ as $N \to \infty$.

Proof of Theorem 1b. Consistency of the Weighted GM Estimator

The objective function of the weighted GM estimator and its nonstochastic counterpart are given by

$$R_N(\underline{\boldsymbol{\theta}}) = (\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \underline{\mathbf{b}})' \boldsymbol{\Theta}_N (\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \underline{\mathbf{b}}) \text{ and} \tag{B.7a}$$

$$\overline{R}_N(\underline{\boldsymbol{\theta}}) = (\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \underline{\mathbf{b}})' \boldsymbol{\Theta}_N (\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \underline{\mathbf{b}}) \tag{B.7b}$$

First, in order to ensure identifiable uniqueness, we show that Assumption 5 also implies that the smallest eigenvalue of $\boldsymbol{\Gamma}'_N \boldsymbol{\Theta}_N \boldsymbol{\Gamma}_N$ is bounded away from zero, i.e.,

$$\lambda_{\min}(\boldsymbol{\Gamma}'_N \boldsymbol{\Theta}_N \boldsymbol{\Gamma}_N) \geq \lambda_0 \text{ for some } \lambda_0 > 0. \tag{B.8}$$

Let $\mathbf{A} = (a_{ij}) = \boldsymbol{\Gamma}_N^{0\prime} \boldsymbol{\Gamma}_N^0$ and $\mathbf{B} = (b_{ij}) = \boldsymbol{\Gamma}_N^{1\prime} \boldsymbol{\Gamma}_N^1$. Note that $\boldsymbol{\Gamma}_N^0$ and $\boldsymbol{\Gamma}_N^1$ are of dimension $(2S+1) \times [2S + S(S-1)/2 + 1]$ (i.e., they have half the rows and one column less than than Γ_N). $\mathbf{A}$ and $\mathbf{B}$ are of order $[2S + S(S-1)/2 + 1] \times [2S + S(S-1)/2 + 1]$ (i.e., they have one row and column less than $\boldsymbol{\Gamma}'_N \boldsymbol{\Gamma}_N$).

Next define $\breve{\boldsymbol{\Gamma}}_N = (\breve{\boldsymbol{\Gamma}}_N^{0\prime}, \breve{\boldsymbol{\Gamma}}_N^{1\prime})'$, which differs from $\boldsymbol{\Gamma}_N$ only by the ordering of the rows. $\breve{\boldsymbol{\Gamma}}_N^0$ corresponds to $\boldsymbol{\Gamma}_N^0$ with a zero column appended as last column, i.e., $\breve{\boldsymbol{\Gamma}}_N^0 = (\boldsymbol{\Gamma}_N^0, 0)$, such that

$$\breve{\boldsymbol{\Gamma}}_N^{0\prime} \breve{\boldsymbol{\Gamma}}_N^0 = \begin{bmatrix} \breve{\boldsymbol{\Gamma}}_N^{0\prime} \breve{\boldsymbol{\Gamma}}_N^0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,2S+S(S-1)/2+1} & 0 \\ \vdots & & & 0 \\ a_{2S+S(S-1)/2+1,1} & \cdots & a_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{B.9a}$$

($\breve{\boldsymbol{\Gamma}}_N^{0\prime} \breve{\boldsymbol{\Gamma}}_N^0$ is of the same dimension as $\boldsymbol{\Gamma}'_N \boldsymbol{\Gamma}_N$, i.e., $[2S + S(S-1)/2 + 2] \times [2S + S(S-1)/2 + 2]$.)

$\breve{\boldsymbol{\Gamma}}_N^1$ is a modified version of $\boldsymbol{\Gamma}_N^1$, with a zero column included as second last column, such that

$$\breve{\boldsymbol{\Gamma}}_N^{1\prime} \breve{\boldsymbol{\Gamma}}_N^1 = \begin{bmatrix} b_{1,1} & \cdots & 0 & b_{1,2S+S(S-1)/2+1} \\ \vdots & & 0 & \vdots \\ 0 & \cdots & 0 & 0 \\ b_{2S+S(S-1)/2+1,1} & \cdots & 0 & b_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} \end{bmatrix}. \tag{B.9b}$$

($\breve{\boldsymbol{\Gamma}}_N^{1\prime} \breve{\boldsymbol{\Gamma}}_N^1$ is of the same dimension as $\boldsymbol{\Gamma}'_N \boldsymbol{\Gamma}_N$, i.e., $[2S + S(S-1)/2 + 2] \times [2S + S(S-1)/2 + 2]$.)

Since $\breve{\boldsymbol{\Gamma}}_N = (\breve{\boldsymbol{\Gamma}}_N^{0\prime}, \breve{\boldsymbol{\Gamma}}_N^{1\prime})'$ differs from $\boldsymbol{\Gamma}_N$ only by the ordering of the rows, it follows that

$$\boldsymbol{\Gamma}'_N \boldsymbol{\Gamma}_N = \breve{\boldsymbol{\Gamma}}'_N \breve{\boldsymbol{\Gamma}}_N = \begin{bmatrix} \breve{\boldsymbol{\Gamma}}_N^{0\prime} & \breve{\boldsymbol{\Gamma}}_N^{1\prime} \end{bmatrix} \begin{bmatrix} \breve{\boldsymbol{\Gamma}}_N^0 \\ \breve{\boldsymbol{\Gamma}}_N^1 \end{bmatrix} = \breve{\boldsymbol{\Gamma}}_N^{0\prime} \breve{\boldsymbol{\Gamma}}_N^0 + \breve{\boldsymbol{\Gamma}}_N^{1\prime} \breve{\boldsymbol{\Gamma}}_N^1, \text{ i.e.,} \tag{B.10}$$

$$\boldsymbol{\Gamma}_N' \boldsymbol{\Gamma}_N = \begin{bmatrix} a_{1,1} & & a_{1,2S+S(S-1)/2+1} & 0 \\ . & & & 0 \\ a_{2S+S(S-1)/2+1,1} & & a_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} b_{1,1} & & 0 & b_{1,2S+S(S-1)/2+1} \\ . & & 0 & . \\ 0 & 0 & 0 & 0 \\ b_{2S+S(S-1)/2+1,1} & . & 0 & b_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} \end{bmatrix}.$$

We can thus write

$$\mathbf{x}'\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N\mathbf{x} = \mathbf{x}'\breve{\boldsymbol{\Gamma}}_N^{0\prime}\breve{\boldsymbol{\Gamma}}_N^{0}\mathbf{x} + \mathbf{x}'\breve{\boldsymbol{\Gamma}}_N^{1\prime}\breve{\boldsymbol{\Gamma}}_N^{1}\mathbf{x} = \mathbf{x}_A'\mathbf{A}_N\mathbf{x}_A + \mathbf{x}_B'\mathbf{B}_N\mathbf{x}_B\,. \tag{B.11}$$

The vector $\mathbf{x}$ is of dimension $[2S + S(S-1)/2 + 2]\times 1$ (corresponding to the number of columns of $\boldsymbol{\Gamma}_N$), wheras $\mathbf{x}_A$ and $\mathbf{x}_B$ are of dimension $[2S + S(S-1)/2 + 1]$, i.e. both have one row less: $\mathbf{x}_A$ excludes the last element of $\mathbf{x}$, i.e., $\mathbf{x}_{2S+S(S-1)/2+2}$, $\mathbf{x}_B$ excludes the second-last element of $\mathbf{x}$, i.e., $\mathbf{x}_{2S+S(S-1)/2+1}$.

Again, we invoke Rao (1973, p. 62) for each quadratic form. It follows

$$\mathbf{x}_A'\mathbf{A}_N\mathbf{x}_A + \mathbf{x}_B'\mathbf{B}_N\mathbf{x}_B \geq \lambda_{\min}(\mathbf{A}_N)\mathbf{x}_A'\mathbf{x}_A + \lambda_{\min}(\mathbf{B}_N)\mathbf{x}_B'\mathbf{x}_B \geq \lambda^*(\mathbf{x}_A'\mathbf{x}_A + \mathbf{x}_B'\mathbf{x}_B) \geq \lambda^*\mathbf{x}'\mathbf{x} \tag{B.12}$$

for any $\mathbf{x} = (x_1, x_2, ..., x_{2S+2})$. Hence, we have shown that $\mathbf{x}'\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N\mathbf{x} \geq \lambda^*\mathbf{x}'\mathbf{x}$, or, equivalently,

$$\frac{\mathbf{x}'\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N\mathbf{x}}{\mathbf{x}'\mathbf{x}} \geq \lambda^* \quad \text{for } \mathbf{x} \neq \mathbf{0}\,. \tag{B.13}$$

Next, note that in light of Rao (1973, p. 62),

$$\lambda_{\min}(\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N) = \inf_x \frac{\mathbf{x}'\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N\mathbf{x}}{\mathbf{x}'\mathbf{x}} \geq \lambda_* > 0\,. \tag{B.14}$$

Using Mittelhammer (1996, p. 254) we have

$$\lambda_{\min}(\boldsymbol{\Gamma}_N'\boldsymbol{\Theta}_N\boldsymbol{\Gamma}_N) = \inf_x \frac{\mathbf{x}'\boldsymbol{\Gamma}_N'\boldsymbol{\Theta}_N\boldsymbol{\Gamma}_N x}{\mathbf{x}'\mathbf{x}} \geq \lambda_{\min}(\Xi_N^{-1})\inf_x \frac{\mathbf{x}'\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N\mathbf{x}}{\mathbf{x}'\mathbf{x}}$$
$$= \lambda_{\min}(\boldsymbol{\Theta}_N)\lambda_{\min}(\boldsymbol{\Gamma}_N'\boldsymbol{\Gamma}_N) \geq \lambda_0 > 0\,, \tag{B.15}$$

with $\lambda_0 = \bar{\lambda}_*\lambda_*$ since $= \lambda_{\min}(\boldsymbol{\Theta}_N) \geq \bar{\lambda}_* > 0$ by assumption (see Theorem 2).

This ensures that the true parameter vector $\boldsymbol{\theta}_N = (\rho_{1,N}, ..., \rho_{S,N}, \sigma_v^2, \sigma_1^2)'$ is identifiable unique.

Next note that in light of the assumptions in Theorem 2, $\boldsymbol{\Theta}_N$ is O(1) by the equivalence of matrix norms. Analogous to the prove of Theorem 1, observe that $\overline{R}_N(\boldsymbol{\theta}) = 0$, i.e., $\overline{R}_N(\underline{\boldsymbol{\theta}}) = 0$ at the true parameter vector $\boldsymbol{\theta}_N = (\rho_{1,N},...,\rho_{S,N},\sigma_v^2,\sigma_1^2)'$. It follows that

$$\overline{R}_N(\underline{\boldsymbol{\theta}}) - \overline{R}_N(\boldsymbol{\theta}_N) = (\underline{\mathfrak{b}} - \mathfrak{b}_N)'\boldsymbol{\Gamma}_N'\boldsymbol{\Theta}_N\boldsymbol{\Gamma}_N(\underline{\mathfrak{b}} - \mathfrak{b}_N). \tag{B.16}$$

Moreover, let $\mathbf{F}_N = (\tilde{\boldsymbol{\gamma}}_N, -\tilde{\boldsymbol{\Gamma}}_N)$ and $\boldsymbol{\Phi}_N = (\boldsymbol{\gamma}_N, -\boldsymbol{\Gamma}_N)$, then,

$$R_N(\underline{\boldsymbol{\theta}}) = (1,\underline{\mathfrak{b}}')\mathbf{F}_N'\boldsymbol{\Theta}_N\mathbf{F}_N(1,\underline{\mathfrak{b}}')' \text{ and} \tag{B.17a}$$

$$\overline{R}_N(\underline{\boldsymbol{\theta}}) = (1,\underline{\mathfrak{b}}')\boldsymbol{\Phi}_N'\boldsymbol{\Theta}_N\boldsymbol{\Phi}_N(1,\underline{\mathfrak{b}}')'. \tag{B.17b}$$

The remainder of the proof is now analogous to that of Theorem 1a.

Proof of Theorem 2. Asymptotic Normality of $\tilde{\boldsymbol{\theta}}_N$

To derive the asymptotic distribution of the vector $\mathbf{q}_N$, defined in (30) we invoke the central limit theorem for vectors of linear quadratic forms given by Kelejian and Prucha (2008, Theorem A.1), which is an extension of the central limit theorem for a single linear quadratic form by Kelejian and Prucha (2001, Theorem 1). The vector of quadratic forms in the present context, to which the Theorem is applied is $\mathbf{q}_N^*$. The variance-covariance matrix of $\mathbf{q}_N$ was derived above and is denoted as $\boldsymbol{\Psi}_N$. Accordingly, the variance-covariance matrix of $\mathbf{q}_N^* = N^{1/2}\mathbf{q}_N$ is given by $\boldsymbol{\Psi}_N^* = N\boldsymbol{\Psi}_N$ and $(\boldsymbol{\Psi}_N^*)^{-1/2} = N^{-1/2}\boldsymbol{\Psi}_N^{-1/2}$.

Note that in light of Assumptions 1, 2 and 7 (and Lemma B.1), the stacked innovations $\boldsymbol{\xi}_N$, the matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s = 1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, and the vectors $\mathbf{a}_{1,s,N}$, …, $\mathbf{a}_{4,s,N}$, $s = 1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$ satisfy the assumptions of central limit theorem by Kelejian and Prucha (2008, Theorem A.1). In the application of Theorem A.1, note that the sample size is given by $NT + N = N(T+1)$ rather than N. As Kelejian and Prucha (2001, p. 227, fn. 13) point out, Theorem A.1 "also holds if the sample size is taken to be k_n rather than n (with $k_n \uparrow \infty$ as $N \to \infty$)." In the present case we have $K_N = (T+1)N$, with $T \geq 1$ and fixed, which ensures that $K_n \uparrow \infty$ as $N \to \infty$. Consequently, Theorem A.1 still applies to each quadratic form in $\mathbf{q}_N^*$. Moreover, as can be observed from the proof of Theorem A.1 in Kelejian and Prucha (2008), the extension of the Theorem from a scalar to a vector of vector of quadratic forms holds up under by this alternative definition of the sample size.

It follows that

$$-(\boldsymbol{\Psi}_N^*)^{-1/2}\mathbf{q}_N^* = -N^{-1/2}\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N^* = -\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} (\mathbf{0},\mathbf{I}_{4S+2}), \tag{B.18}$$

since $N^{-1}\lambda_{\min}(\boldsymbol{\Psi}_N^*) = N^{-1}\lambda_{\min}(N\boldsymbol{\Psi}_N) = \lambda_{\min}(\boldsymbol{\Psi}_N) > 0$ by assumption as required in Theorem A.1.

Since the row and column sums of the matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s=1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, and the vectors $\mathbf{a}_{1,s,N}$, …, $\mathbf{a}_{4,s,N}$, $s=1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$, and the variances $\sigma_{v,N}^2$ and $\sigma_{\mu,N}^2$ are bounded uniformly in absolute value, it follows in light of (38) that the elements of $\boldsymbol{\Psi}_N$ and also those of $\boldsymbol{\Psi}_N^{1/2}$ are bounded uniformly in absolute value.

We next turn to the derivation of the limiting distribution of the GM estimator $\tilde{\boldsymbol{\theta}}_N$. In Theorem 1 we showed that the GM estimator $\tilde{\boldsymbol{\theta}}_N$ defined by (18) is consistent. It follows that – apart from a set of the sample space whose probability tends to zero – the estimator satisfies the following first order condition:[22]

$$\frac{\partial}{\partial \boldsymbol{\rho}}\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)'\tilde{\boldsymbol{\Theta}}_N\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N) = \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'}\tilde{\boldsymbol{\Theta}}_N\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N) = \mathbf{0}, \tag{B.19}$$

which is a $(S+2)\times 1$ vector, the rows corresponding the partial derivatives of the criterion function with respect to $\rho_{s,N}$, $s=1,...,S$, σ_v^2, and σ_1^2.

Substituting the mean value theorem expression

$$\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N) = \mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N) + \frac{\partial \mathbf{q}_N(\bar{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N), \tag{B.20}$$

where $\bar{\boldsymbol{\theta}}_N$ is some between value, into the first-order condition yields

$$\frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\Theta}}_N\frac{\partial \mathbf{q}_N(\bar{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'}N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = -\frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\Theta}}_N N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N). \tag{B.21}$$

Observe that $\frac{\partial \mathbf{q}_N(\boldsymbol{\theta},\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'} = \tilde{\boldsymbol{\Gamma}}_N\mathfrak{B}_N$ and consider the two $(S+2)\times(S+2)$ matrices

$$\tilde{\boldsymbol{\Xi}}_N = \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\Theta}}_N\frac{\partial \mathbf{q}_N(\bar{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'} = \tilde{\mathfrak{B}}_N'\tilde{\boldsymbol{\Gamma}}_N'\tilde{\boldsymbol{\Theta}}_N\tilde{\boldsymbol{\Gamma}}_N\bar{\mathfrak{B}}_N, \tag{B.22}$$

$$\boldsymbol{\Xi}_N = \mathfrak{B}_N'\boldsymbol{\Gamma}_N'\boldsymbol{\Theta}_N\boldsymbol{\Gamma}_N\mathfrak{B}_N, \tag{B.23}$$

where $\tilde{\mathfrak{B}}_N$ and $\bar{\mathfrak{B}}_N$ correspond to $\mathfrak{B}_N$ as defined above with $\tilde{\boldsymbol{\theta}}_N$ and $\bar{\boldsymbol{\theta}}_N$ substituted for $\boldsymbol{\theta}_N$. Notice that $\boldsymbol{\Xi}_N$ is positive definite, since $\boldsymbol{\Gamma}_N$ and $\boldsymbol{\Theta}_N$ are positive definite by assumption and the $[2S+S(S-1)/2+2]\times(S+2)$ matrix $\mathfrak{B}_N$ has full column rank.

[22] The leading two and the negative sign are ignored without further consequences for the proof.

In the proof of Theorem 1 (and Lemma B.1) we have demonstrated that $\widetilde{\boldsymbol{\Gamma}}_N - \boldsymbol{\Gamma}_N \xrightarrow{p} \mathbf{0}$ and that the elements of $\boldsymbol{\Gamma}_N$ and $\widetilde{\boldsymbol{\Gamma}}_N$ are $O(1)$ and $O_p(1)$, respectively. By Assumption 5, $\widetilde{\boldsymbol{\Theta}}_N - \boldsymbol{\Theta}_N = o_p(1)$, $\boldsymbol{\Theta}_N = O(1)$ and $\widetilde{\boldsymbol{\Theta}}_N = O_p(1)$. Since $\widetilde{\boldsymbol{\rho}}_N$ and $\overline{\boldsymbol{\rho}}_N$ (and thus also $\widetilde{\mathfrak{B}}_N$ and $\overline{\mathfrak{B}}_N$) are consistent and bounded uniformly in probability, if follows that $\widetilde{\boldsymbol{\Xi}}_N - \boldsymbol{\Xi}_N = o_p(1)$, $\widetilde{\boldsymbol{\Xi}}_N = O_p(1)$, and $\boldsymbol{\Xi}_N = O(1)$. Moreover, $\boldsymbol{\Xi}_N$ is positive definite and thus invertible, and its inverse $\boldsymbol{\Xi}_N^{-1}$ is also $O(1)$.

Denote $\widetilde{\boldsymbol{\Xi}}_N^{+}$ as the generalized inverse of $\widetilde{\boldsymbol{\Xi}}_N$. It then follows as a special case of Lemma F1 in Pötscher and Prucha (1997) that $\widetilde{\boldsymbol{\Xi}}_N$ is non-singular with probability approaching 1 as $N \to \infty$, that $\widetilde{\boldsymbol{\Xi}}_N^{+}$ is $O_p(1)$, and that $\widetilde{\boldsymbol{\Xi}}_N^{+} - \boldsymbol{\Xi}_N^{-1} = o_p(1)$.

Pre-multiplying (B.21) with $\widetilde{\boldsymbol{\Xi}}_N^{+}$ we obtain, after rearranging terms,

$$N^{1/2}(\widetilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = (\mathbf{I}_{S+2} - \widetilde{\boldsymbol{\Xi}}_N^{+}\widetilde{\boldsymbol{\Xi}}_N)N^{1/2}(\widetilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) - N^{1/2}\widetilde{\boldsymbol{\Xi}}_N^{+}\frac{\partial \mathbf{q}_N(\widetilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\widetilde{\boldsymbol{\Theta}}_N \mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N). \quad \text{(B.24)}$$

In light of the discussion above, the first term on the right-hand side is zero on ω-sets of probability approaching 1 (compare Pötscher and Prucha, 1997, pp. 228). This yields

$$N^{1/2}(\widetilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = -\widetilde{\boldsymbol{\Xi}}_N^{+}\frac{\partial \mathbf{q}_N(\widetilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\widetilde{\boldsymbol{\Theta}}_N N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N) + o_p(1). \quad \text{(B.25)}$$

Next observe that

$$\widetilde{\boldsymbol{\Xi}}_N^{+}\frac{\partial \mathbf{q}_N(\widetilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\widetilde{\boldsymbol{\Theta}}_N - \boldsymbol{\Xi}_N^{-1}\mathfrak{B}'_N\boldsymbol{\Gamma}'_N\boldsymbol{\Theta}_N = o_p(1), \quad \text{(B.26)}$$

since $\widetilde{\boldsymbol{\Xi}}_N^{+} - \boldsymbol{\Xi}_N^{-1} = o_p(1)$ and $\frac{\partial \mathbf{q}_N(\widetilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}} - \mathfrak{B}'_N\boldsymbol{\Gamma}'_N = o_p(1)$.

As we showed in section III, the elements of $N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N)$ can be expressed as

$$N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N) = N^{-1/2}\mathbf{q}_N^{*} + o_p(1) = \mathbf{q}_N + o_p(1). \quad \text{(B.27)}$$

where $\mathbf{q}_N^{*}$ is defined in (27), and that

$$-(\boldsymbol{\Psi}_N^{*})^{-1/2}\mathbf{q}_N^{*} = -N^{-1/2}\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N^{*} = -\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} (\mathbf{0}, \mathbf{I}_{4S+2}). \quad \text{(B.28)}$$

It now follows from (B.25), (B.26), and (B.27) that

$$N^{1/2}(\widetilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = \boldsymbol{\Xi}_N^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N^{1/2}(-\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N) + o_p(1). \quad \text{(B.29)}$$

Since all nonstochastic terms on the right hand side from (B.29) are $O(1)$ it follows that $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ is $O_p(1)$. To derive the asymptotic distribution of $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$, we invoke Corollary F4 in Pötscher and Prucha (1997). In the present context, we have

$$\boldsymbol{\zeta}_N = -\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} \boldsymbol{\zeta} \sim N(\mathbf{0}, \mathbf{I}_{4S+2}),$$
$$N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = \boldsymbol{\mathfrak{A}}_N\boldsymbol{\zeta}_N + o_p(1) \text{ , with}$$
$$\boldsymbol{\mathfrak{A}}_N = \boldsymbol{\Xi}_N^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N^{1/2}.$$

Furthermore, $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = O_p(1)$ and its variance-covariance matrix is

$$\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\boldsymbol{\Theta}_N) = (\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N\boldsymbol{\Theta}_N\mathbf{J}_N(\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1},$$

where $\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}$ is positive definite.

As a final point it has to be shown that $\liminf_{N\to\infty} \lambda_{\min}(\boldsymbol{\mathfrak{A}}_N\boldsymbol{\mathfrak{A}}'_N) > 0$ as required in Corollary F4 in Pötscher and Prucha (1997). Observe that

$$\begin{aligned}\lambda_{\min}(\boldsymbol{\mathfrak{A}}_N\boldsymbol{\mathfrak{A}}'_N) &= \lambda_{\min}(\boldsymbol{\Xi}_N^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N\boldsymbol{\Theta}'_N\mathbf{J}_N\boldsymbol{\Xi}_N^{-1}) \\ &\geq \lambda_{\min}(\boldsymbol{\Psi}_N)\lambda_{\min}(\boldsymbol{\Theta}_N\boldsymbol{\Theta}'_N)\lambda_{\min}(\boldsymbol{\Xi}_N^{-1}\boldsymbol{\Xi}_N^{-1})\lambda_{\min}(\boldsymbol{\Gamma}'_N\boldsymbol{\Gamma}_N)\lambda_{\min}(\boldsymbol{\mathfrak{B}}'_N\boldsymbol{\mathfrak{B}}_N) > 0,\end{aligned} \tag{B.30}$$

since the matrices involved are all positive definite.

Consistency Proof for Estimates of Third and Fourth Moments of the Error Components

Consistent estimates for the second moments of $v_{it,N}$ and $\mu_{i,N}$ are delivered by the GM estimators defined in (17) and (18), respectively (See Theorems 1a and 1b). In the technical appendix, it is shown that the estimators for the third and fourth moments of $v_{it,N}$ and $\mu_{it,N}$, defined in (39) and (40) are also consistent.

III. Proof of Theorem 3 (Variance-Covariance Estimation)

Lemma B.2

Suppose Assumptions 1-4 hold. Furthermore, assume that $\sup_N \sum_{m=1}^{S} |\rho_{m,N}| < 1$, and that the row and column sums of $\mathbf{M}_{m,N}$, $m = 1,...,S$ are bounded uniformly in absolute value by 1 and some finite constant respectively. Let $\tilde{\sigma}_v^2$, $\tilde{\sigma}_1^2$, and $\tilde{\rho}_{s,N}$, $s = 1,...,S$, be estimators, satisfying $\tilde{\sigma}_v^2 - \sigma_v^2 = o_p(1)$, $\tilde{\sigma}_1^2 - \sigma_1^2 = o_p(1)$, $\tilde{\rho}_{s,N} - \rho_{s,N} = o_p(1)$, $s = 1,...,S$. Let the $NT \times NT$ or $N \times N$ matrix $\mathbf{F}_N$ be of the form (compare Lemmata 1 and 2):

(a) $\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$,

$\mathbf{F}_{\mu,N} = (\mathbf{e}'_T \otimes \mathbf{I}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$,

(b) $\mathbf{F}^{**}_{v,N} = (\sigma_v^{-2}\mathbf{Q}_{0,N} + \sigma_1^{-2}\mathbf{Q}_{1,N})\mathbf{H}^{*}_N = \mathbf{\Omega}^{-1}_{\boldsymbol{\varepsilon},N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}_{m,N})]\mathbf{H}_N$,

$\mathbf{F}^{**}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)]\mathbf{H}^{*}_N = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}_{m,N})]\mathbf{H}_N$,

where $\mathbf{H}_N$ is a $N \times P_*$ matrix whose elements are bounded uniformly in absolute value by some constant $c < \infty$. The corresponding estimates $\widetilde{\mathbf{F}}_{v,N}$, $\widetilde{\mathbf{F}}_{\mu,N}$, $\widetilde{\mathbf{F}}^{**}_{v,N}$, and $\widetilde{\mathbf{F}}^{**}_{\mu,N}$ are defined analogously, replacing σ_v^2, σ_1^2, $\mathbf{\Omega}_{\boldsymbol{\varepsilon},N}$, and $\rho_{m,N}$, $m = 1,...,S$, with $\widetilde{\sigma}_v^2, \widetilde{\sigma}_1^2$, $\widetilde{\mathbf{\Omega}}_{\boldsymbol{\varepsilon},N}$, and $\widetilde{\rho}_{m,N}$, $m = 1,...,S$, respectively.

(i) Then, $N^{-1}\widetilde{\mathbf{F}}'_N\widetilde{\mathbf{F}}_N - N^{-1}\mathbf{F}'_N\mathbf{F}_N = o_p(1)$ and $N^{-1}\mathbf{F}'_N\mathbf{F}_N = O(1)$.

(ii) Let $\mathbf{a}_N$ be some $NT \times 1$ or $N \times 1$ vector, whose elements are bounded uniformly in absolute value. Then $N^{-1}\widetilde{\mathbf{F}}'_N\mathbf{a}_N - N^{-1}\mathbf{F}'_N\mathbf{a}_N = o_p(1)$ and $N^{-1}\mathbf{F}'_{v,N}\mathbf{a}_N = O(1)$.

The proof of Lemma B.2 is given in the technical appendix.

Proof of Theorem 3

As part of proving Theorem 3 it has to be shown that $\widetilde{\mathbf{\Psi}}_N - \mathbf{\Psi}_N = o_p(1)$. Observe that in light of (15), each element of $\widetilde{\mathbf{\Psi}}_N$ and the corresponding element of $\mathbf{\Psi}_N$ can be written as

$$\widetilde{\mathfrak{E}}^{p,q}_{r,s,N} = \widetilde{\mathfrak{E}}^{p,q,*}_{r,s,N} + \widetilde{\mathfrak{E}}^{p,q,**}_{r,s,N} + \widetilde{\mathfrak{E}}^{p,q,***}_{r,s,N} + \widetilde{\mathfrak{E}}^{p,q,****}_{r,s,N},$$
$$\mathfrak{E}^{p,q}_{r,s,N} = \mathfrak{E}^{p,q,*}_{r,s,N} + \mathfrak{E}^{p,q,**}_{r,s,N} + \mathfrak{E}^{p,q,***}_{r,s,N} + \mathfrak{E}^{p,q,****}_{r,s,N},$$

for $p,q = 1,...,4,(a,b)$, $r,s = 1,...,S+1$[23], where

$$\widetilde{\mathfrak{E}}^{p,q,*}_{r,s,N} = 2N^{-1}\widetilde{\sigma}^4_{v,N} Tr(\mathbf{A}^{v}_{p,r,N}\mathbf{A}^{v}_{q,s,N}) + 2N^{-1}\widetilde{\sigma}^4_{\mu,N} Tr(\mathbf{A}^{\mu}_{p,r,N}\mathbf{A}^{\mu}_{q,s,N}),$$
$$\widetilde{\mathfrak{E}}^{p,q,**}_{r,s,N} = \widetilde{\sigma}^2_{v,N} N^{-1}\widetilde{\mathbf{a}}^{v\,\prime}_{p,r,N}\,\widetilde{\mathbf{a}}^{v}_{q,s,N} + \widetilde{\sigma}^2_{\mu,N} N^{-1}\widetilde{\mathbf{a}}^{\mu\,\prime}_{p,r,N}\,\widetilde{\mathbf{a}}^{\mu}_{q,s,N}$$
$$\widetilde{\mathfrak{E}}^{p,q,***}_{r,s,N} = \widetilde{\sigma}^{(3)}_{v,N} N^{-1}\sum_{i=1}^{NT}(\widetilde{a}^{v}_{p,r,i,N}a^{v}_{q,s,ii,N} + a^{v}_{p,r,ii,N}\widetilde{a}^{v}_{q,s,i,N}) + \widetilde{\sigma}^{(3)}_{\mu,N} N^{-1}\sum_{i=1}^{N}(\widetilde{a}^{\mu}_{p,r,i,N}a^{\mu}_{q,s,ii,N} + \widetilde{a}^{\mu}_{p,r,ii,N}a^{\mu}_{q,s,i,N}),$$
$$\widetilde{\mathfrak{E}}^{p,q,****}_{r,s,N} = (\widetilde{\sigma}^{(4)}_{v,N} - 3\widetilde{\sigma}^4_{v,N})N^{-1}\sum_{i=1}^{NT} a^{v}_{p,r,ii,N}\widetilde{a}^{v}_{q,s,ii,N} + (\widetilde{\sigma}^{(4)}_{\mu,N} - 3\widetilde{\sigma}^4_{\mu,N})N^{-1}\sum_{i=1}^{N}\widetilde{a}^{\mu}_{p,r,ii,N}a^{\mu}_{q,s,ii,N},$$

and

[23] See equations (34) in the main text for the structure of the matrix $\mathbf{\Psi}_N$ and the proper indexation of its elements.

$$\mathfrak{E}^{p,q,*}_{r,s,N} = 2N^{-1}\sigma_v^4 Tr(\mathbf{A}^v_{p,r,N}\mathbf{A}^v_{q,s,N}) + 2N^{-1}\sigma_\mu^4 Tr(\mathbf{A}^\mu_{p,r,N}\mathbf{A}^\mu_{q,s,N}),$$

$$\mathfrak{E}^{p,q,**}_{r,s,N} = \sigma_v^2 N^{-1}\mathbf{a}^{v\,\prime}_{p,r,N}\,\mathbf{a}^v_{q,s,N} + \sigma_\mu^2 N^{-1}\mathbf{a}^{\mu\,\prime}_{p,r,N}\,\mathbf{a}^\mu_{q,s,N},$$

$$\mathfrak{E}^{p,q,***}_{r,s,N} = \sigma_v^{(3)} N^{-1}\sum_{i=1}^{NT}(a^v_{p,r,i,N}a^v_{q,s,ii,N} + a^v_{p,r,ii,N}a^v_{q,s,i,N}) + \sigma_\mu^{(3)} N^{-1}\sum_{i=1}^{N}(a^\mu_{p,r,i,N}a^\mu_{q,s,ii,N} + a^\mu_{p,r,ii,N}a^\mu_{q,s,i,N}),$$

$$\mathfrak{E}^{p,q,****}_{r,s,N} = (\sigma_v^{(4)} - 3\sigma_v^4)N^{-1}\sum_{i=1}^{NT} a^v_{p,r,ii,N}a^v_{q,s,ii,N} + (\sigma_\mu^{(4)} - 3\sigma_\mu^4)N^{-1}\sum_{i=1}^{N} a^\mu_{p,r,ii,N}a^\mu_{q,s,ii,N}.$$

Invoking Lemma B.2, it is shown in the technical appendix that $\tilde{\mathfrak{E}}^{p,q,*}_{r,s,N} - \mathfrak{E}^{p,q,*}_{r,s,N} = o_p(1)$, $\tilde{\mathfrak{E}}^{p,q,**}_{r,s,N} - \mathfrak{E}^{p,q,**}_{r,s,N} = o_p(1)$, $\tilde{\mathfrak{E}}^{p,q,***}_{r,s,N} - \mathfrak{E}^{p,q,***}_{r,s,N} = o_p(1)$, and $\tilde{\mathfrak{E}}^{p,q,****}_{r,s,N} - \mathfrak{E}^{p,q,****}_{r,s,N} = o_p(1)$ for all elements of $\mathbf{\Psi}_N$, $p,q = 1,...,4,(a,b)$, $r,s = 1,...,S+1$. It follows that $\tilde{\mathbf{\Psi}}_N - \mathbf{\Psi}_N = o_p(1)$.

Notice further that $\tilde{\mathbf{\Theta}}_N - \mathbf{\Theta}_N = o_p(1)$, $\mathbf{\Theta}_N = O(1)$ and $\tilde{\mathbf{\Theta}}_N = O_p(1)$ by Assumption 5. Let $\mathbf{\Xi}_N = \mathbf{J}'_N\mathbf{\Theta}\mathbf{J}_N = \mathfrak{B}'_N\,\mathbf{\Gamma}'_N\,\mathbf{\Theta}_N\mathbf{\Gamma}_N\mathfrak{B}_N$ (as in Theorem 2) and $\tilde{\mathbf{\Xi}}_N = \tilde{\mathbf{J}}'_N\tilde{\mathbf{\Theta}}\tilde{\mathbf{J}}_N = \tilde{\mathfrak{B}}'_N\,\tilde{\mathbf{\Gamma}}'_N\tilde{\mathbf{\Theta}}_N\tilde{\mathbf{\Gamma}}_N\tilde{\mathfrak{B}}_N$.[24] In Theorem 2, we showed that $\tilde{\mathbf{J}}_N = O_p(1)$, $\mathbf{J}_N = O(1)$, and $\tilde{\mathbf{J}}_N - \mathbf{J}_N = o_p(1)$ and that $\tilde{\mathbf{\Xi}}^+_N = O_p(1)$, $\mathbf{\Xi}^{-1}_N = O_p(1)$ and $\tilde{\mathbf{\Xi}}^+_N - \mathbf{\Xi}^{-1}_N = o_p(1)$. It now follows that $\tilde{\mathbf{\Omega}}_{\tilde{\boldsymbol{\theta}}_N} - \mathbf{\Omega}_{\tilde{\boldsymbol{\theta}}_N} = o_p(1)$.

III. Proof of Theorem 4 (Joint Distribution of $\tilde{\boldsymbol{\rho}}_N$ and Other Model Parameters)

The subsequent proof will focus on the case $\mathbf{F}_{v,N}$ and $\mathbf{F}_{v,N}$; this also convers the case for $\mathbf{F}^{**}_{v,N}$ and $\mathbf{F}^{**}_{\mu,N}$. The first line in Theorem 4 holds in light of Assumption 7 (for $N^{1/2}\mathbf{\Delta}_N$), bearing in mind that $\mathbf{T}_{v,N} = \mathbf{F}_{v,N}\mathbf{P}_N$, and Theorem 2 (for $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$).

We next prove that $\boldsymbol{\xi}_{o,N} = \mathbf{\Psi}^{-1/2}_{o,N}(N^{-1/2}\boldsymbol{\xi}'_N\mathbf{F}_N, \mathbf{q}'_N)' \xrightarrow{d} N(\mathbf{0}, \mathbf{I}_{P^*+4S+2})$ by verifying that the assumptions of the central limit theorem A.1 by Kelejian and Prucha (2008) are fulfilled. Note that $\lambda_{\min}(\mathbf{\Psi}_{o,N}) \geq c^*_{\mathbf{\Psi}_o} > 0$ by assumption. In Theorem 2, we verified that the stacked innovations $\boldsymbol{\xi}_N$, the matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s = 1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, and the vectors $\mathbf{a}_{1,s,N}$, …, $\mathbf{a}_{4,s,N}$, $s = 1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$ satisfy the assumptions of central limit theorem by Kelejian and Prucha (2008, Theorem A.1).

Next, consider the two blocks of $\mathbf{F}_N = (\mathbf{F}'_{v,N}, \mathbf{F}'_{\mu,N})'$, which are given by

[24] There is a slight discrepancy to the definition of $\tilde{\mathbf{\Xi}}_N$ in Theorem 2: Here $\tilde{\mathfrak{B}}_N$ is used rather than $\overline{\mathfrak{B}}_N$, which does not affect the proof, however, noting that both $\tilde{\boldsymbol{\rho}}_N$ and $\overline{\boldsymbol{\rho}}_N$ are consistent.

$$\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N \text{, and}$$

$$\mathbf{F}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)]\mathbf{\Omega}_{\boldsymbol{\varepsilon},N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N .$$

Since the row and columns sums of $[\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)]$, $\mathbf{\Omega}_{\boldsymbol{\varepsilon},N}$, and $[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}'_{m,N})^{-1}]$ are uniformly bounded in absolute value and since the elements of the matrix $\mathbf{H}_N$ are uniformly bounded in absolute value, it follows that the elements of $\mathbf{F}_N$ are also uniformly bounded in absolute value. Hence, the linear form $\mathbf{F}'_N\boldsymbol{\xi}_N = \mathbf{F}'_{v,N}\mathbf{v}_N + \mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N$ also fulfils the assumptions of Theorem A.1. As a consequence, $\boldsymbol{\xi}_{o,N} \xrightarrow{d} N(\mathbf{0}, \mathbf{I}_{P^*+4S+2})$.

In the proofs of Theorems 2 and 3, we showed that $\tilde{\mathbf{\Psi}}_N - \mathbf{\Psi}_N = o_p(1)$, $\mathbf{\Psi}_N = O(1)$, and $\tilde{\mathbf{\Psi}}_N = O_p(1)$. By analogous arguments, this also holds for the submatrices $\mathbf{\Psi}_{\Delta\Delta,N}$ and $\mathbf{\Psi}_{\Delta\theta,N}$. Hence, $\tilde{\mathbf{\Psi}}_{o,N} - \mathbf{\Psi}_{o,N} = o_p(1)$, $\mathbf{\Psi}_{o,N} = O(1)$ and $\tilde{\mathbf{\Psi}}_{o,N} - \mathbf{\Psi}_{o,N} = o_p(1)$, and thus $\tilde{\mathbf{\Psi}}_{o,N} = O_p(1)$.

By assumption $\tilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$, $\mathbf{P}_N = O(1)$, and $\tilde{\mathbf{P}}_N = O_p(1)$ as well as $\tilde{\mathbf{\Theta}}_N - \mathbf{\Theta}_N = o_p(1)$, $\mathbf{\Theta}_N = O(1)$ and $\tilde{\mathbf{\Theta}}_N = O_p(1)$. In the proof of Theorem 2 we showed that $\tilde{\mathbf{J}}_N - \mathbf{J}_N = o_p(1)$, $\mathbf{J}_N = O(1)$, and $\tilde{\mathbf{J}}_N = O_p(1)$, and that $(\tilde{\mathbf{J}}'_N\tilde{\mathbf{\Theta}}_N\tilde{\mathbf{J}}_N)^+ - (\mathbf{J}'_N\mathbf{\Theta}_N\mathbf{J}_N)^{-1} = o_p(1)$, $(\mathbf{J}'_N\mathbf{\Theta}_N\mathbf{J}_N)^{-1} = O(1)$, and $(\tilde{\mathbf{J}}'_N\tilde{\mathbf{\Theta}}_N\tilde{\mathbf{J}}_N)^+ = O_p(1)$. It now follows that $\tilde{\mathbf{\Omega}}_{o,N} - \mathbf{\Omega}_{o,N} = o_p(1)$ and $\mathbf{\Omega}_{o,N} = O(1)$ and thus $\tilde{\mathbf{\Omega}}_{o,N} = O_p(1)$.

APPENDIX C.

Proof of Lemma 1.

In light of equations (4a) and (4b), Assumptions 3 and 8, as well as $\sup_N \|\boldsymbol{\beta}_N\| \leq b < \infty$, it follows that all columns of $\mathbf{Z}_N = (\mathbf{X}_N, \overline{\mathbf{Y}}_N)$ are of the form $\vartheta_N = \boldsymbol{\pi}_N + \mathbf{\Pi}_N\boldsymbol{\varepsilon}_N$, where the elements of the vector $\boldsymbol{\pi}_N$ and the row and column sums of the matrix $\mathbf{\Pi}_N$ are bounded uniformly in absolute value (see Remark A.1 in Appendix A). It follows from Lemma C.2 in Kelejian and Prucha (2008) that the fourth moments of the elements of the matrix $\mathbf{D}_N = -\mathbf{Z}_N$ are bounded uniformly by some finite constant and that Assumption 6 holds.

Next, note that

$$(NT)^{1/2}(\tilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N) = \tilde{\mathbf{P}}'_N(NT)^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N + \tilde{\mathbf{P}}'_N(NT)^{-1/2}\mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N ,$$

where $\tilde{\mathbf{P}}_N$ is defined in the Lemma, and

$$\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N \text{ and}$$

$$\mathbf{F}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)]\mathbf{\Omega}_{\varepsilon,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N .$$

In light of Assumption 10, $\widetilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$ and $\mathbf{P}_N = O(1)$, with $\mathbf{P}_N$ as defined in the Lemma. By Assumptions 2, 3 and 9, the elements of $\mathbf{F}_{v,N}$ and $\mathbf{F}_{\mu,N}$ are bounded uniformly in absolute value. By Assumption 2, $E(\mathbf{v}_N) = \mathbf{0}$, $E(\boldsymbol{\mu}_N) = \mathbf{0}$, and the diagonal variance-covariance matrices of $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ have uniformly bounded elements. Thus, $E[(NT)^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N] = \mathbf{0}$ and the elements of the variance-covariance matrix of $N^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N$, i.e., $(NT)^{-1}\sigma_v^2\mathbf{F}'_{v,N}\mathbf{F}_{v,N}$, are bounded uniformly in absolute value (see Remark A.1 in Appendix A). Moreover, $E[(NT)^{-1/2}\mathbf{F}'_{2,N}\boldsymbol{\mu}_N] = \mathbf{0}$, and the elements of the variance-covariance matrix of $N^{-1/2}\mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N$, i.e., $(NT)^{-1}\sigma_\mu^2\mathbf{F}'_{\mu,N}\mathbf{F}_{\mu,N}$, are bounded uniformly in absolute value. It follows from Chebychev's inequality that $(NT)^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N = O_p(1)$, $(NT)^{-1/2}\mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N = O_p(1)$, and consequently $(NT)^{1/2}(\widetilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N) = \mathbf{P}_N(NT)^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N + \mathbf{P}_N(NT)^{-1/2}\mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N + o_p(1)$ and $\mathbf{P}_N(NT)^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N + \mathbf{P}_N(NT)^{-1/2}\mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N = O_p(1)$. This completes the proof, recalling that $\mathbf{T}_N = (\mathbf{T}'_{v,N}, \mathbf{T}'_{\mu,N})' = (\mathbf{P}'_N\mathbf{F}'_{v,N}, \mathbf{P}'_N\mathbf{F}'_{\mu,N})'$.

Proof of Lemma 2.

The structure of the proof of Lemma 2 is similar to that of Lemma 1, applied to the transformed model and accounting for the use of generated instruments. It is given in full detail in the technical appendix.

Table 1. Parameter Constellations in Monte Carlo Experiments

Parameter constellation	λ_1	λ_2	λ_3	ρ_1	ρ_2	ρ_3
(1)	0.5	0.3	0.1	0.4	0.25	0.1
(2)	0.5	0	0	0.4	0	0
(3)	0	0	0	0	0	0

Note: $\beta_1 = \beta_2 = 1$ under all parameter constellations.

Table 2. Monte Carlo Results

Parameter constellation[1)]	N = 100			N = 500		
	(1)	(2)	(3)	(1)	(2)	(3)
λ_1	0.5	0.4	0	0.5	0.4	0
Bias	0.0004	0.0014	0.0013	0.0005	0.0000	0.0003
RMSE	0.0203	0.0230	0.0244	0.0088	0.0100	0.0099
Rej. Rate	0.0540	0.0590	0.0490	0.0590	0.0710	0.0380
λ_2	0.3	0	0	0.3	0	0
Bias	0.0008	0.0001	0.0001	-0.0002	0.0000	-0.0001
RMSE	0.0213	0.0226	0.0251	0.0094	0.0097	0.0104
Rej. Rate	0.0490	0.0520	0.0620	0.0620	0.0410	0.0480
λ_3	0.1	0	0	0.1	0	0
Bias	-0.0003	-0.0005	0.0010	0.0001	-0.0002	0.0000
RMSE	0.0213	0.0232	0.0250	0.0093	0.0102	0.0101
Rej. Rate	0.0520	0.0490	0.0690	0.0630	0.0530	0.0490
β_1	1	1	1	1	1	1
Bias	0.0001	-0.0004	-0.0003	0.0001	0.0000	0.0000
RMSE	0.0134	0.0132	0.0138	0.0061	0.0060	0.0061
Rej. Rate	0.0560	0.0500	0.0560	0.0550	0.0600	0.0480
β_2	1	1	1	1	1	1
Bias	-0.0007	-0.0002	-0.0001	0.0001	-0.0001	0.0002
RMSE	0.0130	0.0142	0.0133	0.0060	0.0058	0.0059
Rej. Rate	0.0460	0.0740	0.0550	0.0500	0.0520	0.0510
ρ_1	0.4	0.3	0	0.4	0.3	0
Bias	-0.0050	-0.0064	-0.0073	0.0013	0.0025	0.0027
RMSE	0.0946	0.1037	0.1261	0.0385	0.0426	0.0496
Rej. Rate	0.1070	0.1200	0.1330	0.0890	0.0910	0.0940
ρ_2	0.25	0	0	0.25	0	0
Bias	-0.0091	-0.0036	-0.0047	-0.0007	0.0002	0.0008
RMSE	0.1077	0.1107	0.1214	0.0444	0.0433	0.0477
Rej. Rate	0.1180	0.1090	0.1140	0.0870	0.0810	0.0790
ρ_3	0.1	0	0	0.1	0	0
Bias	-0.0079	-0.0020	-0.0028	-0.0027	0.0002	-0.0003
RMSE	0.1005	0.1044	0.1169	0.0404	0.0423	0.0475
Rej. Rate	0.0900	0.0980	0.0920	0.0790	0.0780	0.0860
Joint Tests [2)]						
Rej. Rate	-	0.1280	0.1510	-	0.0790	0.0880

Note: [1)] Each column corresponds to one parameter constellation (see Table 1). [2)] Rejections rates for the following hypotheses: (2): $H_0^{\lambda,\rho,*}: \lambda_2 = \lambda_3 = \rho_2 = \rho_3 = 0$; (3): $H_0^{\lambda,\rho}: \lambda_1 = \lambda_2 = \lambda_3 = \rho_1 = \rho_2 = \rho_3 = 0$.

TECHNICAL APPENDIX

The proofs for the panel data error component SARAR(*R*,*S*) framework are given in full length to the benefit of the reader. They build on analogous proofs by Kelejian and Prucha (2008) for a cross-sectional SARAR(1,1) and Badinger and Egger (2008b) for a cross-section SARAR(*R*,*S*) model, as well as analogous proofs for a panel SARAR(0,1) model with nonstochastic regressors by Kapoor, Kelejian, and Prucha (2007).

APPENDIX A

Notation

We adopt the standard convention to refer to matrices and vectors with acronyms in boldface. Let $\mathbf{A}_N$ denote some matrix. Its elements are referred to as $a_{ij,N}$; $\mathbf{a}_{i.,N}$ and $\mathbf{a}_{.i,N}$ denote the *i*-th row and the *i*-th column of $\mathbf{A}_N$ respectively. If $\mathbf{A}_N$ is a square matrix, $\mathbf{A}_N^{-1}$ denotes its inverse; if $\mathbf{A}_N$ is singular, $\mathbf{A}_N^{+}$ denotes its generalized inverse. If $\mathbf{A}_N$ is a square, symmetric and positive definite matrix, $\mathbf{A}_N^{1/2}$ denotes the unique positive definite square root of $\mathbf{A}_N$ and $\mathbf{A}_N^{-1/2}$ denotes $(\mathbf{A}_N^{-1})^{1/2}$. The (submultiplicative) matrix norm $\|\cdot\|$ is defined as $\|\mathbf{A}_N\| = [Tr(\mathbf{A}_N'\mathbf{A}_N)]^{1/2}$. Finally, unless stated otherwise, for expressions involving sums over elements of vectors or matrices that are stacked over all time periods, we adopt the convention to use single indexation *i*, running from $i = 1,...,NT$, to denote elements of the stacked vectors or matrices.[1]

Remark A.1

i) Definition of row and column sum boundedness (compare Kapoor, Kelejian, and Prucha, 2007, p. 99): Let $\mathbf{B}_N, N \geq 1$, be some sequence of $NT \times NT$ matrices with T some fixed positive integer. We will then say that the row and column sums of the (sequence of) matrices $\mathbf{B}_N$ are bounded uniformly in absolute value, if there exists a constant $c < \infty$, which does not depend on *N*, such that

$$\max_{1 \leq i \leq NT} \sum_{j=1}^{NT} |b_{ij,N}| \leq c \text{ and } \max_{1 \leq j \leq NT} \sum_{i=1}^{NT} |b_{ij,N}| \leq c \text{ for all } N \geq 1.$$

The following results will be repeatedly used in the subsequent proofs.

[1] Take the vector $\mathbf{u}_N = [\mathbf{u}_N'(1),...,\mathbf{u}_N'(T)]$, for example. Using indexation $i = 1,...,NT$, the elements $u_{i,N}, i = 1,...,N$ refer to $t = 1$, elements $u_{i,N}, i = N+1,...,2N$ refer to $t = 2$, etc., and elements $u_{i,N}, i = (T-1)N+1,...,NT$ refer to $t = T$. The major advantage of this notation is that it avoids the use of double indexation for the cross-section and time dimension. Moreover, it allows us the invoke several results referring to the case of a single cross-section, which still apply to the case of T stacked cross-sections.

ii) Let $\mathbf{R}_N$ be a (sequence of) $N \times N$ matrices whose row and column sums are bounded uniformly in absolute value, and let $\mathbf{S}$ be some $T \times T$ matrix (with $T \geq 1$ fixed). Then the row and column sums of the matrix $\mathbf{S} \otimes \mathbf{R}_N$ are bounded uniformly in absolute value (compare Kapoor, Kelejian, and Prucha, 2007, p. 118).

iii) If $\mathbf{A}_N$ and $\mathbf{B}_N$ are (sequences of) $NT \times NT$ matrices (with $T \geq 1$ fixed), whose row and column sums are bounded uniformly in absolute value, then so are the row and column sums of $\mathbf{A}_N\mathbf{B}_N$ and $\mathbf{A}_N + \mathbf{B}_N$. If $\mathbf{Z}_N$ is a (sequence of) $NT \times P$ matrices whose elements are bounded uniformly in absolute value, then so are the elements of $\mathbf{A}_N\mathbf{Z}_N$ and $(NT)^{-1}\mathbf{Z}_N'\mathbf{A}_N\mathbf{Z}_N$. Of course, this also covers the case $(NT)^{-1}\mathbf{Z}_N'\mathbf{Z}_N$ for $\mathbf{A}_N = \mathbf{I}_{NT}$ (compare Kapoor, Kelejian, and Prucha, 2007, p. 119).

iv) Suppose that the row and columns sums of the $NT \times NT$ matrices $\mathbf{A}_N = (a_{ij,N})$ are bounded uniformly in absolute value by some finite constant c_A; then $\sum_{i=1}^{NT} \left| a_{ij,N} \right|^q \leq c_A^q$ for $q > 1$ (see Kelejian and Prucha, 2008, Remark C.1).

v) Let $\boldsymbol{\xi}_N$ and $\boldsymbol{\eta}_N$ be $NT \times 1$ random vectors (with $T \geq 1$ fixed), where, for each *N*, the elements are independently distributed with zero mean and finite variances. Then the elements of $(NT)^{-1/2}\mathbf{Z}_N'\boldsymbol{\xi}_N$ are $O_p(1)$ and $(NT)^{-1}\boldsymbol{\xi}_N'\mathbf{A}_N\boldsymbol{\eta}_N$ is $O_p(1)$ (compare Kelejian and Prucha, 2004, Remark A.1).[2]

vi) Let $\boldsymbol{\zeta}_N$ be a $NT \times 1$ random vector (with $T \geq 1$ fixed), where, for each *N*, the elements are distributed with zero mean and finite fourth moments. Let $\boldsymbol{\pi}_N$ be some nonstochastic $NT \times 1$ vector, whose elements are bounded uniformly in absolute value and let $\boldsymbol{\Pi}_N$ be a $NT \times NT$ nonstochastic matrix whose row and column sums are bounded uniformly in absolute value. Define the column vector $\mathbf{d}_N = \boldsymbol{\pi}_N + \boldsymbol{\Pi}_N\boldsymbol{\zeta}_N$. It follows that the elements of $\mathbf{d}_N$ have finite fourth moments. (Compare Kelejian and Prucha, 2008, Lemma C.2, for the case $T = 1$ and independent elements of $\boldsymbol{\zeta}_N$.)[3]

---

[2] Kelejian and Prucha (2004) consider the case $T = 1$ and where the elements of $\boldsymbol{\xi}_N$ and $\boldsymbol{\eta}_N$ are identically distributed. Obviously, the results also holds for (fixed) $T \geq 1$ and under heteroskedasticity, as long as the variances of the elements of $\boldsymbol{\xi}_N$ and $\boldsymbol{\eta}_N$ are bounded uniformly in absolute value.

[3] The extension to (fixed) $T \geq 1$ is obvious. Independence of the elements of $\boldsymbol{\zeta}_N$ is not required for the result to hold. The fourth moments of the elements of $\mathbf{d}_N = \boldsymbol{\pi}_N + \boldsymbol{\Pi}_N\boldsymbol{\zeta}_N$ are given by $E(\pi_{i,N} + \sum_{j=1}^{NT} \pi_{ij,N}\zeta_{j,N})^4 \leq 2^4 E[\pi_{i,N}^4 + (\sum_{j=1}^{NT} \pi_{ij,N}\zeta_{j,N})^4]$

Remark A2.
The matrices $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ have the following properties (see Kapoor, Kelejian, and Prucha, 2007, p. 101):

$$tr(\mathbf{Q}_{0,N}) = N(T-1),\ tr(\mathbf{Q}_{1,N}) = N,\ \mathbf{Q}_{0,N}(\mathbf{e}_T \otimes \mathbf{I}_N) = \mathbf{0},\ \mathbf{Q}_{1,N}(\mathbf{e}_T \otimes \mathbf{I}_N) = (\mathbf{e}_T \otimes \mathbf{I}_N),$$
$$\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N = \mathbf{Q}_{0,N}\mathbf{v}_N,\ \mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N = (\mathbf{e}_T \otimes \mathbf{I}_N)\boldsymbol{\mu}_N + \mathbf{Q}_{1,N}\mathbf{v}_N,\ (\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{0,N} = \mathbf{Q}_{0,N}(\mathbf{I}_T \otimes \mathbf{D}_N),$$
$$(\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{1,N} = \mathbf{Q}_{1,N}(\mathbf{I}_T \otimes \mathbf{D}_N),\ tr[(\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{0,N}] = (T-1)tr(\mathbf{D}_N),$$
$$tr[(\mathbf{I}_T \otimes \mathbf{D}_N)\mathbf{Q}_{1,N}] = tr(\mathbf{D}_N),$$

where $\mathbf{D}_N$ is an arbitrary $N \times N$ matrix. Obviously, the row and column sums of $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$ are bounded uniformly in absolute value.

APPENDIX B

The following lemma will be repeatedly used in the subsequent proofs.

Lemma B.1[4]

Let $\mathbf{A}_N$ be some nonstochastic $NT \times NT$ matrix (with T fixed), whose row and column sums are bounded uniformly in absolute value. Let $\mathbf{u}_N$ be defined by (2c) and $\tilde{\mathbf{u}}_N$ be a predictor for $\mathbf{u}_N$. Suppose that Assumptions 1 to 4 hold. Then

(a) $N^{-1}E\left|\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N\right| = O(1)$, $Var(N^{-1}\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N) = o(1)$,
and $N^{-1}(\tilde{\mathbf{u}}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N) - N^{-1}E(\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N) = o_p(1)$.

(b) $N^{-1}E\left|\mathbf{d}'_{.j,N}\mathbf{A}_N\mathbf{u}_N\right| = O(1)$, $j = 1,...,P$, where $\mathbf{d}_{.j,N}$ is the j-th column of the $NT \times P$ matrix $\mathbf{D}_N$, and $N^{-1}\mathbf{D}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N - N^{-1}E(\mathbf{D}'_N\mathbf{A}_N\mathbf{u}_N) = o_p(1)$.

(c) If furthermore Assumption 6 holds, then
$N^{-1/2}\tilde{\mathbf{u}}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N = N^{-1/2}\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N + \boldsymbol{\alpha}'_N N^{1/2}\boldsymbol{\Delta}_N + o_p(1)$ with $\boldsymbol{\alpha}_N = N^{-1}E[\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\mathbf{u}_N]$.

$\leq 2^4[\pi^4_{i,N} + \sum_{j=1}^{NT}\sum_{k=1}^{NT}\sum_{l=1}^{NT}\sum_{m=1}^{NT}\left|\pi_{ij,N}\right\|\pi_{ik,N}\left\|\pi_{il,N}\right\|\pi_{im,N}\left|E\right|\zeta_{j,N}\left\|\zeta_{k,N}\right\|\zeta_{l,N}\left\|\zeta_{m,N}\right|] \leq K < \infty$, by Hölder's inequality as long as the fourth moments of the elements of $\boldsymbol{\zeta}_N$ are bounded uniformly.

[4] Compare Lemma C.1 in Kelejian and Prucha (2008) for the case of a cross-sectional SARAR(1,1) model and Lemma C.1 in Badinger and Egger (2008b) for the case of a cross-sectional SARAR(*R*,*S*) model.

In light of (b), we have $\boldsymbol{\alpha}_N = O(1)$ and $N^{-1}\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\tilde{\mathbf{u}}_N - \boldsymbol{\alpha}_N = o_p(1)$.

Proof of part (a)

Let

$$\vartheta_N = N^{-1}\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N \text{ and } \tilde{\vartheta}_N = N^{-1}\tilde{\mathbf{u}}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N, \tag{B.1}$$

then given (4a), we have $\vartheta_N = N^{-1}\boldsymbol{\varepsilon}'_N\mathbf{B}_N\boldsymbol{\varepsilon}_N$, with the symmetric $NT \times NT$ matrix $\mathbf{B}_N$ defined as

$$\mathbf{B}_N = (1/2)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})^{-1}](\mathbf{A}_N + \mathbf{A}'_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}] \tag{B.2}$$

By Assumptions 1-3 and Remark A.1 in Appendix A, the row and column sums of the matrices $\mathbf{B}_N$ and $\boldsymbol{\Omega}_{\varepsilon,N}$ are bounded uniformly in absolute value. It follows that the row and column sums of the matrices $\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}$ are bounded uniformly in absolute value.

In the following let $K < \infty$ be a common bound for the row and column sums of the absolute value of the elements of $\mathbf{B}_N, \boldsymbol{\Omega}_{\varepsilon,N}$, and $\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}$ and of the absolute value of their respective elements. Then

$$\begin{aligned} E|\vartheta_N| &= E\left|N^{-1}\sum_{i=1}^{NT}\sum_{j=1}^{NT} b_{ij,N}\varepsilon_{i,N}\varepsilon_{j,N}\right| \qquad \text{(B.3)} \\ &\le N^{-1}\sum_{i=1}^{NT}\sum_{j=1}^{NT}|b_{ij,N}|E|\varepsilon_{i,N}\|\varepsilon_{j,N}| \\ &\le N^{-1}\sum_{i=1}^{NT}\sum_{j=1}^{NT}|b_{ij,N}|\sigma_{\varepsilon,i}\sigma_{\varepsilon,j} \\ &\le TK^3, \end{aligned}$$

where we used Hölder's inequality in the last step. This proves that $E|\vartheta_N|$ is O(1).

Now consider $Var(\vartheta_N)$, invoking Lemma A.1 in KP (2008):

$$\begin{aligned} Var(\vartheta_N) &= Cov(N^{-1}\boldsymbol{\varepsilon}'_N\mathbf{B}_N\boldsymbol{\varepsilon}_N, N^{-1}\boldsymbol{\varepsilon}'_N\mathbf{B}_N\boldsymbol{\varepsilon}_N) \qquad \text{(B.4a)} \\ &= 2N^{-2}Tr(\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}) + N^{-2}\sum_{i=1}^{NT} b_{ii,*}^2[\mu_{\eta_i,N}^{(4)} - 3], \\ &= 2N^{-2}Tr(\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}\mathbf{B}_N\boldsymbol{\Omega}_{\varepsilon,N}) + N^{-2}Tr\{diag_{i=1,...,NT}(b_{ii*,N}^2)diag_{i=1,...,NT}[E(\eta_{i,N}^4) - 3]\} \qquad \text{(B.4b)} \end{aligned}$$

where $b_{ii*,N}$ is the *i*-th diagonal element of $\mathbf{B}_N^* = (b_{ij,*,N}) = \mathbf{S}_N' \mathbf{B}_N' \mathbf{S}_N$ and $\mu_{\eta_i,N}^{(4)}$ is the fourth moment of the *i*-th element of the $NT \times 1$ vector $\boldsymbol{\eta}_N = \mathbf{S}_N^{-1} \boldsymbol{\varepsilon}_N$, i.e., $\mu_{\eta_i,N}^{(4)} = E(\eta_{i,N}^4)$. In light of Assumption 1 and the properties and $\mathbf{Q}_{0,N}$ and $\mathbf{Q}_{1,N}$, the row and column sums (and the elements) of $\mathbf{S}_N = \sigma_v \mathbf{Q}_{0,N} + \sigma_1 \mathbf{Q}_{1,N}$ are bounded uniformly in absolute value by some finite constant, say K^*. W.o.l.o.g. we can choose the bound K used above such that $K^* \leq K$.

Moreover, the row and column sums (and the elements) of $\mathbf{S}_N^{-1} = \sigma_v^{-1} \mathbf{Q}_{0,N} + \sigma_1^{-1} \mathbf{Q}_{1,N}$ are also bounded uniformly in absolute value by some constant K^{**}. W.o.l.o.g. we can choose K such that $K^{**} \leq K$.

In light of Remark A.1 and Assumption 1 it follows that the elements of $\boldsymbol{\eta}_N = \mathbf{S}^{-1} \boldsymbol{\varepsilon}_N$ have finite fourth moments. Denote their bound by K^{***}. W.o.l.o.g. we assume that $K^{***} \leq K$.

Hence, we have

$$
\begin{aligned}
&Var(\vartheta_N) \leq 2N^{-2} Tr(K \mathbf{I}_{NT}) + N^{-2} Tr[diag_{i=1,...,NT}(K^2 K)]. \\
&= 2N^{-1} TK + N^{-1} TK^3 = N^{-1}(2TK + TK^3) = o(1).
\end{aligned}
$$

The claim in part (a) of Lemma B.1 that $N^{-1}(\mathbf{u}_N' \mathbf{A}_N \mathbf{u}_N) - N^{-1} E(\mathbf{u}_N' \mathbf{A}_N \mathbf{u}_N) = o_p(1)$ now follows from Chebychev's inequality (see, for example, White, 2001, p. 35).

We now prove the second part of (a), i.e., $N^{-1}(\tilde{\mathbf{u}}_N' \mathbf{A}_N \tilde{\mathbf{u}}_N) - N^{-1} E(\mathbf{u}_N' \mathbf{A}_N \mathbf{u}_N) = o_p(1)$. Since $\vartheta_N - E(\vartheta_N) = o_p(1)$, it suffices to show that $\tilde{\vartheta}_N - \vartheta_N = o_p(1)$. By Assumption 4, we have $\tilde{\mathbf{u}}_N - \mathbf{u}_N = \mathbf{D}_N \boldsymbol{\Delta}_N$, where $\mathbf{D}_N = (\mathbf{d}_{1.,N}', ..., \mathbf{d}_{NT.,N}')'$. Substituting $\tilde{\mathbf{u}}_N = \mathbf{u}_N + \mathbf{D}_N \Delta_N$ into the expression for $\tilde{\vartheta}_N$ in (B.1), we obtain

$$
\begin{aligned}
&\tilde{\vartheta}_N - \vartheta_N = N^{-1}(\mathbf{u}_N' + \boldsymbol{\Delta}_N' \mathbf{D}_N') \mathbf{A}_N (\mathbf{u}_N + \mathbf{D}_N \boldsymbol{\Delta}_N) - N^{-1} \mathbf{u}_N' \mathbf{A}_N \mathbf{u}_N \qquad \text{(B.5)} \\
&= N^{-1}(\mathbf{u}_N' \mathbf{A}_N \mathbf{u}_N + \boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{u}_N + \mathbf{u}_N' \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N + \boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N - \mathbf{u}_N' \mathbf{A}_N \mathbf{u}_N) \\
&= N^{-1}(\boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{u}_N + \mathbf{u}_N' \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N + \boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N) \\
&= N^{-1}(\boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{u}_N + \boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N' \mathbf{u}_N + \boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N) \\
&= N^{-1}[\boldsymbol{\Delta}_N' \mathbf{D}_N' (\mathbf{A}_N + \mathbf{A}_N') \mathbf{u}_N + \boldsymbol{\Delta}_N' \mathbf{D}_N' \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N] \\
&= \phi_N + \psi_N,
\end{aligned}
$$

where

$$\phi_N = N^{-1}[\boldsymbol{\Delta}'_N \mathbf{D}'_N (\mathbf{A}_N + \mathbf{A}'_N)\mathbf{u}_N] = N^{-1}\{\boldsymbol{\Delta}'_N \mathbf{D}'_N (\mathbf{A}_N + \mathbf{A}'_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N\}, \quad \text{(B.6)}$$

$$= N^{-1}(\boldsymbol{\Delta}'_N \mathbf{D}'_N \mathbf{C}_N \boldsymbol{\varepsilon}_N), \text{ with}$$

$$\mathbf{C}_N = (\mathbf{A}_N + \mathbf{A}'_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N})^{-1}] = (\mathbf{c}'_{1.,N},..,\mathbf{c}'_{NT.,N})',$$

$$\psi_N = N^{-1}\boldsymbol{\Delta}'_N \mathbf{D}'_N \mathbf{A}_N \mathbf{D}_N \boldsymbol{\Delta}_N. \quad \text{(B.7)}$$

By Assumption 3 and Remark A.1, the row and column sums of $\mathbf{C}_N$ are bounded uniformly in absolute value. We next prove that $\phi_N = o_p(1)$ and $\psi_N = o_p(1)$.

Proof that $\phi_N = o_p(1)$:

$$|\phi_N| = N^{-1}|\boldsymbol{\Delta}'_N \mathbf{D}'_N \mathbf{C}_N \boldsymbol{\varepsilon}_N| \quad \text{(B.8)}$$

$$= N^{-1}\left\|\sum_{i=1}^{NT} \boldsymbol{\Delta}'_N \mathbf{d}'_{i.,N} \mathbf{c}_{i.,N} \boldsymbol{\varepsilon}_N\right\|$$

$$\leq N^{-1}\sum_{i=1}^{NT} \|\boldsymbol{\Delta}'_N\| \|\mathbf{d}'_{i.,N}\| \|\mathbf{c}_{i.,N}\boldsymbol{\varepsilon}_N\|$$

$$= N^{-1}\|\boldsymbol{\Delta}'_N\| \sum_{i=1}^{NT} \|\mathbf{d}'_{i.,N}\| \left\|\sum_{j=1}^{NT} c_{ij,N}\varepsilon_{j,N}\right\|$$

$$\leq N^{-1}\|\boldsymbol{\Delta}'_N\| \sum_{i=1}^{NT} \|\mathbf{d}'_{i.,N}\| \sum_{j=1}^{NT} |c_{ij,N}\varepsilon_{j,N}|$$

$$= N^{-1}\|\boldsymbol{\Delta}'_N\| \sum_{i=1}^{NT} \|\mathbf{d}'_{i.,N}\| \sum_{j=1}^{NT} |c_{ij,N}| |\varepsilon_{j,N}|$$

$$= N^{-1}\|\boldsymbol{\Delta}'_N\| \sum_{j=1}^{NT} |\varepsilon_{j,N}| \sum_{i=1}^{NT} \|\mathbf{d}'_{i.,N}\| |c_{ij,N}|$$

$$\leq N^{-1}\|\boldsymbol{\Delta}'_N\| \sum_{j=1}^{NT} |\varepsilon_{j,N}| \left(\sum_{i=1}^{NT} \|\mathbf{d}'_{i.,N}\|^p\right)^{1/p} \left(\sum_{i=1}^{NT} |c_{ij,N}|^q\right)^{1/q}$$

$$= N^{-1}N^{-1/2}NN^{1/p}\left(N^{1/2}\|\boldsymbol{\Delta}'_N\|\right)\left(N^{-1}\sum_{j=1}^{NT}|\varepsilon_{j,N}|\right)\left(N^{-1}\sum_{i=1}^{NT}\|\mathbf{d}'_{i.,N}\|^p\right)^{1/p}\left(\sum_{i=1}^{NT}|c_{ij,N}|^q\right)^{1/q}$$

$$= N^{1/p-1/2}\left(N^{1/2}\|\boldsymbol{\Delta}_N\|\right)\left(N^{-1}\sum_{j=1}^{NT}|\varepsilon_{j,N}|\right)\left(N^{-1}\sum_{i=1}^{NT}\|\mathbf{d}'_{i.,N}\|^p\right)^{1/p}\left(\sum_{i=1}^{NT}|c_{ij,N}|^q\right)^{1/q}$$

Note that $\left(\sum_{i=1}^{NT} |c_{ij,N}|\right) \leq \overline{K} < \infty$ by Assumption. In the following we denote by $\overline{K}$ the uniform bound for the row and column sums of the absolute value of the elements of $\mathbf{A}_N$ and $\mathbf{C}_N$.

From Remark C.1 in KP (2008) (see Remark A.1 in Appendix A) it follows that $\left(\sum_{i=1}^{NT}\left|c_{ij,N}\right|^{q}\right) \leq \overline{K}^{q}$ and thus $\left(\sum_{i=1}^{NT}\left|c_{ij,N}\right|^{q}\right)^{1/q} \leq \overline{K}$. Factoring $\overline{K}$ out of the sum yields

$$\left|\phi_N\right| \leq \overline{K} N^{1/p-1/2}\left(N^{1/2}\left\|\boldsymbol{\Delta}_N\right\|\right)\left(N^{-1}\sum_{j=1}^{NT}\left|\varepsilon_{j,N}\right|\right)\left(T(NT)^{-1}\sum_{i=1}^{NT}\left\|\mathbf{d}'_{i.,N}\right\|^{p}\right)^{1/p}.$$

This holds for $p=2+\delta$ for some $\delta>0$ as in Assumption 4 and $1/p+1/q=1$. By Assumption 4, $\left(N^{1/2}\left\|\boldsymbol{\Delta}_N\right\|\right)=O_p(1)$. Assumption 4 also implies that $\left((NT)^{-1}\sum_{i=1}^{NT}\left\|\mathbf{d}'_{i.,N}\right\|^{p}\right)^{1/p}=O_p(1)$ for $p=2+\delta$ and some $\delta>0$.

Moreover, $E\left|\varepsilon_{j,N}\right| \leq K<\infty$, which implies that $\left(N^{-1}\sum_{j=1}^{NT}\left|\varepsilon_{j,N}\right|\right)=O_p(1)$. Since $N^{1/p-1/2}\to 0$ as $N\to\infty$ it follows that $\left|\phi_N\right|=o_p(1)$.

Next consider

$$\begin{aligned}
\left|\psi_N\right| &= N^{-1}\left|\boldsymbol{\Delta}'_N\mathbf{D}'_N\mathbf{A}_N\mathbf{D}_N\boldsymbol{\Delta}_N\right| = N^{-1}\left|\sum_{i=1}^{NT}\sum_{j=1}^{NT}\boldsymbol{\Delta}'_N\mathbf{d}'_{i.,N}a_{ij,N}\mathbf{d}_{j.,N}\boldsymbol{\Delta}_N\right| \qquad \text{(B.9)}\\
&\leq N^{-1}\left\|\boldsymbol{\Delta}_N\right\|^2\sum_{i=1}^{NT}\left\|\mathbf{d}'_{i.,N}\right\|\left\|\sum_{j=1}^{NT}\left\|\mathbf{d}_{j.,N}\right\|\left|a_{ij,N}\right|\right.\\
&\leq N^{-1}\left\|\boldsymbol{\Delta}_N\right\|^2\sum_{i=1}^{NT}\left\|\mathbf{d}'_{i.,N}\right\|\left(\sum_{j=1}^{NT}\left\|\mathbf{d}_{j.,N}\right\|^{p}\right)^{1/p}\left(\sum_{j=1}^{NT}\left|a_{ij,N}\right|^{q}\right)^{1/q}\\
&\leq N^{1/p}\overline{K}\left\|\boldsymbol{\Delta}_N\right\|^2\left(N^{-1}\sum_{i=1}^{NT}\left\|\mathbf{d}'_{i.,N}\right\|\right)\left(N^{-1}\sum_{j=1}^{NT}\left\|\mathbf{d}_{j.,N}\right\|^{p}\right)^{1/p}\\
&= N^{1/p-1/2}N^{-1/2}\overline{K}\left(N^{1/2}\left\|\boldsymbol{\Delta}_N\right\|\right)^2\left(N^{-1}\sum_{j=1}^{NT}\left\|\mathbf{d}_{j.,N}\right\|^{p}\right)^{2/p}=o_p(1).
\end{aligned}$$

From the last inequality we can also see that $N^{1/2}\psi_N=o_p(1)$. Summing up, we have proved that $\tilde{\vartheta}_N-\vartheta_N=\phi_N+\psi_N=o_p(1)$.

Proof of part (b)

Denote by $\vartheta^*_{s,N}$ the *s*-th element of $N^{-1}\mathbf{D}'_N\mathbf{A}_N\mathbf{u}_N$. By Assumptions 3 and 4 and Remark A.1 in Appendix A there exists a constant $\overline{\overline{K}} < \infty$ such that $E(u^2_{i,N}) \le \overline{\overline{K}}$ and $E\left|d_{ij,N}\right|^p \le \overline{\overline{K}}$ with $p = 2+\delta$ for some $\delta > 0$. W.o.l.o.g. we assume that the row and column sums of the matrices $\mathbf{A}_N$ are bounded uniformly by $\overline{\overline{K}} < \infty$. Notice first that

$$
\begin{aligned}
&E\left|u_{i,N}\right\|d_{js,N}\right| \le \left(Eu^2_{i,N}\right)^{1/2}\left(Ed^2_{js,N}\right)^{1/2} \\
&\le \left(Eu^2_{i,N}\right)^{1/2}\left(E\left|d_{js,N}\right|^p\right)^{1/p} \\
&\le \overline{\overline{K}}^{1/2}\overline{\overline{K}}^{1/p} = \overline{\overline{K}}^{1/2+1/p} \text{ with } p \text{ as before.}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&E\left|\vartheta^*_{s,N}\right| = N^{-1}\sum_{i=1}^{NT}\sum_{j=1}^{NT}\left|a_{ij,N}\right| E\left|u_{i,N}\right\|d_{js,N}\right| \qquad \text{(B.10)} \\
&\le \overline{\overline{K}}^{1/2+1/p} N^{-1}\sum_{i=1}^{NT}\sum_{j=1}^{NT}\left|a_{ij,N}\right| \le \overline{\overline{K}}^{1/2+1/p} N^{-1} NT\overline{\overline{K}} = T\overline{\overline{K}}^{3/2+1/p} < \infty,
\end{aligned}
$$

which shows that indeed $E\left|N^{-1}\mathbf{d}'_{.s,N}\mathbf{A}_N\mathbf{u}_N\right| = O(1)$. Of course, the argument also shows that

$$\boldsymbol{\alpha}_N = N^{-1}E[\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\mathbf{u}_N] = O(1).$$

It is readily verified that $Var(\vartheta^*_s) = o(1)$, such that we have $\vartheta^*_s - E(\vartheta^*_s) = o_p(1)$. Next observe that

$$N^{-1}\mathbf{D}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N = N^{-1}\mathbf{D}'_N\mathbf{A}_N\mathbf{u}_N + \phi^*_N, \qquad \text{(B.11)}$$

where $\phi^*_N = N^{-1}\mathbf{D}'_N\mathbf{A}_N\mathbf{D}_N\boldsymbol{\Delta}_N$. By arguments analogous to the proof that $\phi_N = N^{-1}[\boldsymbol{\Delta}'_N\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\mathbf{u}_N] = o_p(1)$, it follows that $\phi^*_N = o_p(1)$. Hence $\tilde{\vartheta}^*_s - \vartheta^*_s = o_p(1)$, and thus $\tilde{\vartheta}^*_s - E(\vartheta^*_s) = o_p(1)$, which also shows that $N^{-1}\mathbf{D}'_N(\mathbf{A}_N + \mathbf{A}'_N)\tilde{\mathbf{u}}_N - \boldsymbol{\alpha}_N = o_p(1)$.

Proof of part (c)

In light of the proof of part (a)

$$N^{-1/2}\tilde{\mathbf{u}}'_N\mathbf{A}_N\tilde{\mathbf{u}}_N = N^{-1/2}\mathbf{u}'_N\mathbf{A}_N\mathbf{u}_N + [N^{-1}\mathbf{u}'_N(\mathbf{A}_N + \mathbf{A}'_N)\mathbf{D}_N]N^{1/2}\boldsymbol{\Delta}_N + N^{1/2}\psi_N, \qquad \text{(B.12)}$$

where $N^{1/2}\psi_N = o_p(1)$ as shown above, and in light of (b) and since $N^{1/2}\boldsymbol{\Delta}_N = O_p(1)$ by Assumption 4, we have

$$N^{-1/2}\tilde{\mathbf{u}}_N'\mathbf{A}_N\tilde{\mathbf{u}}_N = N^{-1/2}\mathbf{u}_N'\mathbf{A}_N\mathbf{u}_N + N^{1/2}\boldsymbol{\alpha}_N'\boldsymbol{\Delta}_N + o_p(1). \tag{B.13}$$

Proof of Theorem 1a. Consistency of Initial GM Estimator $\hat{\boldsymbol{\theta}}_N^0$

The objective function of the nonlinear least squares estimator in (17a) and its nonstochastic counterpart are given by

$$\mathbf{R}_N^0(\omega,\underline{\boldsymbol{\theta}}^0) = (\tilde{\boldsymbol{\gamma}}_N^0 - \tilde{\boldsymbol{\Gamma}}_N^0\underline{\mathfrak{b}}_N^0)'(\tilde{\boldsymbol{\gamma}}_N^0 - \tilde{\boldsymbol{\Gamma}}_N^0\underline{\mathfrak{b}}_N^0) \text{ and} \tag{B.14a}$$

$$\overline{\mathbf{R}}_N^0(\underline{\boldsymbol{\theta}}^0) = (\boldsymbol{\gamma}_N^0 - \boldsymbol{\Gamma}_N^0\underline{\mathfrak{b}}^0)'(\boldsymbol{\gamma}_N^0 - \boldsymbol{\Gamma}_N^0\underline{\mathfrak{b}}^0). \tag{B.14b}$$

Since $\boldsymbol{\gamma}_N^0 - \boldsymbol{\Gamma}_N^0\mathfrak{b}_N^0 = 0$, we have $\overline{R}_N^0(\boldsymbol{\theta}_N^0) = 0$, i.e., $\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) = 0$ at the true parameter vector $\boldsymbol{\theta}_N^0 = (\rho_1,...,\rho_S,\sigma_v^2)'$. Hence,

$$\overline{\mathbf{R}}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{\mathbf{R}}_N^0(\boldsymbol{\theta}^0) = (\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'\boldsymbol{\Gamma}_N^{0\prime}\boldsymbol{\Gamma}_N^0(\underline{\mathfrak{b}}_N^0 - \mathfrak{b}_N^0). \tag{B.15}$$

In light of Rao (1973, p. 62) and Assumption 5, it follows that:

$$\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0) \geq \lambda_{\min}(\boldsymbol{\Gamma}_N^{0\prime}\boldsymbol{\Gamma}_N^0)(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0) \text{ and}$$

$$\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0) \geq \lambda_*(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0).$$

By the properties of the norm $\|\mathbf{A}\| = [tr(\mathbf{AA})]^{1/2}$, we have $\left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\|^2 \leq (\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)'(\underline{\mathfrak{b}}^0 - \mathfrak{b}_N^0)$ such that $\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0) \geq \lambda_*\left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\|^2$. Hence, for every $\varepsilon > 0$

$$\lim_{N\to\infty} \inf_{\{\boldsymbol{\theta}^0 : \left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\| \geq \varepsilon\}} [\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\boldsymbol{\theta}^0)] \geq \inf_{\{\boldsymbol{\theta}^0 : \left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\| \geq \varepsilon\}} \lambda_*\left\|\underline{\boldsymbol{\theta}}^0 - \boldsymbol{\theta}^0\right\|^2 = \lambda_*\varepsilon^2 > 0, \tag{B.16}$$

which proves that the true parameter $\boldsymbol{\theta}^0$ is identifiable unique.

Next, let $\mathbf{F}_N^0 = (\tilde{\boldsymbol{\gamma}}_N^0, -\tilde{\boldsymbol{\Gamma}}_N^0)$ and $\boldsymbol{\Phi}_N^0 = (\boldsymbol{\gamma}_N^0, -\boldsymbol{\Gamma}_N^0)$. The objective function and its nonstochastic counterpart can then be written as

$R_N^0(\underline{\boldsymbol{\theta}}^0) = (1, \underline{\mathbf{b}}^{0\prime})\mathbf{F}_N^{0\prime}\mathbf{F}_N^0(1, \underline{\mathbf{b}}^{0\prime})'$ and

$\overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) = (1, \underline{\mathbf{b}}^{0\prime})\boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0(1, \underline{\mathbf{b}}^{0\prime})'$.

Hence for $\underline{\boldsymbol{\rho}} \in [-\mathbf{a}^\rho, \mathbf{a}^\rho]$ [5] and $\underline{\sigma}_v^2 \in [0, b]$ it holds that

$$\left\| R_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) \right\| = \left\| (1, \underline{\mathbf{b}}^{0\prime})(\mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0)(1, \underline{\mathbf{b}}^{0\prime})' \right\|.$$

Moreover, since the norm $\|\cdot\|$ is submultiplicative, i.e., $\|\mathbf{AB}\| \le \|\mathbf{A}\|\|\mathbf{B}\|$, we have

$$\left\| R_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) \right\| \le \left\| \mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0 \right\| \left\| (1, \underline{\mathbf{b}}_N^{0\prime}) \right\|^2$$
$$\le \left\| \mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0 \right\| [1 + S(a^\rho)^2 + \frac{2S + S(S-1)}{2}(a^\rho)^4 + b^2].$$

It is readily observed from (16), that the elements of the matrices $\boldsymbol{\gamma}_N^0$ and $\boldsymbol{\Gamma}_N^0$ are all of the form $\mathbf{u}_N'\mathbf{A}_N\mathbf{u}_N$, where $\mathbf{A}_N$ are nonstochastic $NT \times NT$ matrices (with T fixed), whose row and column sums are bounded uniformly in absolute value. In light of Lemma B.1, the elements of $\boldsymbol{\Phi}_N^0$ are O(1) and it follows that $\left\| \mathbf{F}_N^0 - \boldsymbol{\Phi}_N^0 \right\| \xrightarrow{p} 0$ and $\left\| \mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0 \right\| \xrightarrow{p} 0$ as $N \to \infty$. As a consequence, we have (for finite S)

$$\sup_{\underline{\boldsymbol{\rho}} \in [-\mathbf{a}^\rho, \mathbf{a}^\rho], \underline{\sigma}_v^2 \in [0,b]} \left| R_N^0(\underline{\boldsymbol{\theta}}^0) - \overline{R}_N^0(\underline{\boldsymbol{\theta}}^0) \right| \le \left\| [\mathbf{F}_N^{0\prime}\mathbf{F}_N^0 - \boldsymbol{\Phi}_N^{0\prime}\boldsymbol{\Phi}_N^0] \right\| [1 + S(a^\rho)^2 + \frac{S(S-1)}{2}(a^\rho)^4 + b^2] \xrightarrow{p} 0 \text{ as } N \to \infty \tag{B.17}$$

Together with identifiable uniqueness, the consistency of $\widetilde{\boldsymbol{\theta}}_N^0 = (\widetilde{\rho}_{1,N}, ..., \widetilde{\rho}_{S,N}, \widetilde{\sigma}_{v,N}^2)$ now follows directly from Lemma 3.1 in Pötscher and Prucha (1997).

Having proved that the estimators $\widetilde{\rho}_{1,N} ..., \widetilde{\rho}_{S,N}, \widetilde{\sigma}_{v,N}^2$ are consistent for $\rho_{1,N} ..., \rho_{S,N}, \sigma_v^2$, we now show that σ_1^2 can be estimated consistently from the last line $(4S+2)$ of equation system (12), using

$$\widetilde{\sigma}_{1,N}^2 = \widetilde{\gamma}_{4S+2,N} - \widetilde{\gamma}_{4S+2,1,N}\widetilde{\rho}_{1,N} - ... - \widetilde{\gamma}_{4S+2,S,N}\widetilde{\rho}_{S,N} - \widetilde{\gamma}_{4S+2,S+1,N}\widetilde{\rho}_{1,N}^2$$
$$- \widetilde{\gamma}_{4S+2,2S,N}\widetilde{\rho}_{S,N}^2 - \widetilde{\gamma}_{4S+2,2S+1,N}\widetilde{\rho}_{1,N}\widetilde{\rho}_{2,N} - ... - \widetilde{\gamma}_{4S+2,2S+S(S-1)/2,N}\widetilde{\rho}_{S-1,N}\widetilde{\rho}_{S,N}. \tag{B.18}$$

[5] This should be read as $\underline{\rho}_s \in [-a^\rho, a^\rho]$ for all $s = 1, ..., S$.

Since $\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \mathfrak{b}_N^0 = 0$, we have

$$
\begin{aligned}
&\tilde{\sigma}_1^2 - \sigma_1^2 = (\tilde{\gamma}_{4S+2,N} - \gamma_{4S+2,N}) - (\tilde{\gamma}_{4S+2,1,N} - \gamma_{4S+2,1,N})\tilde{\rho}_{1,N} - \ldots - (\tilde{\gamma}_{4S+2,S,N} - \gamma_{4S+2,S,N})\tilde{\rho}_{S,N} \\
&- (\tilde{\gamma}_{4S+2,S+1,N} - \gamma_{4S+2,S+1,N})\tilde{\rho}_{1,N}^2 - \ldots - (\tilde{\gamma}_{4S+2,2S,N} - \gamma_{4S+2,2S,N})\tilde{\rho}_{S,N}^2 \\
&- (\tilde{\gamma}_{4S+2,2S+1,N} - \gamma_{4S+2,2S+1,N})\tilde{\rho}_{1,N}\tilde{\rho}_{2,N} - \ldots - (\tilde{\gamma}_{4S+2,2S+S(S-1)/2,N} - \gamma_{4S+2,2S+S(S-1)/2,N})\tilde{\rho}_{S-1,N}\tilde{\rho}_{S,N} \\
&- \gamma_{4S+2,1,N}(\tilde{\rho}_{1,N} - \rho_{1,N}) - \ldots - \gamma_{4S+2,S,N}(\tilde{\rho}_{S,N} - \rho_{S,N}) \\
&- \gamma_{4S+2,S+1,N}(\tilde{\rho}_{1,N}^2 - \rho_{1,N}^2) \ldots - \gamma_{4S+2,2S,N}(\tilde{\rho}_{S,N}^2 - \rho_{S,N}^2) \\
&- \gamma_{4S+2,2S+1,N}(\tilde{\rho}_{1,N}\tilde{\rho}_{2,N}^2 - \rho_{1,N}\rho_{2,N}) \ldots - \gamma_{4S+2,2S+S(S-1)/2,N}(\tilde{\rho}_{S-1,N}\tilde{\rho}_{S,N} - \rho_{S-1,N}\rho_{S,N}).
\end{aligned}
\tag{B.19}
$$

Since $\left\| \mathbf{F}_N^0 - \boldsymbol{\Phi}_N^0 \right\| \xrightarrow{p} 0$ as $N \to \infty$ and the elements of $\boldsymbol{\Phi}_N$ are O(1) it follows from the consistency of $\tilde{\rho}_{1,N}, \ldots, \tilde{\rho}_{S,N}$ that $\tilde{\sigma}_{1,N}^2 - \sigma_1^2 \xrightarrow{p} 0$ as $N \to \infty$.

Proof of Theorem 1b. Consistency of the Weighted GM Estimator

The objective function of the weighted GM estimator and its nonstochastic counterpart are given by

$$R_N(\underline{\boldsymbol{\theta}}) = (\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \underline{\mathfrak{b}})' \boldsymbol{\Theta}_N (\tilde{\boldsymbol{\gamma}}_N - \tilde{\boldsymbol{\Gamma}}_N \underline{\mathfrak{b}}) \text{ and} \tag{B.20a}$$

$$\bar{R}_N(\underline{\boldsymbol{\theta}}) = (\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \underline{\mathfrak{b}})' \boldsymbol{\Theta}_N (\boldsymbol{\gamma}_N - \boldsymbol{\Gamma}_N \underline{\mathfrak{b}}) \tag{B.20b}$$

First, in order to ensure identifiable uniqueness, we show that Assumption 5 also implies that the smallest eigenvalue of $\boldsymbol{\Gamma}_N' \boldsymbol{\Theta}_N \boldsymbol{\Gamma}_N$ is bounded away from zero, i.e.,

$$\lambda_{\min}(\boldsymbol{\Gamma}_N' \boldsymbol{\Theta}_N \boldsymbol{\Gamma}_N) \geq \lambda_0 \text{ for some } \lambda_0 > 0. \tag{B.21}$$

Let $\mathbf{A} = (a_{ij}) = \boldsymbol{\Gamma}_N^{0\prime} \boldsymbol{\Gamma}_N^0$ and $\mathbf{B} = (b_{ij}) = \boldsymbol{\Gamma}_N^{1\prime} \boldsymbol{\Gamma}_N^1$. Note that $\boldsymbol{\Gamma}_N^0$ and $\boldsymbol{\Gamma}_N^1$ are of dimension $(2S+1) \times [2S + S(S-1)/2 + 1]$ (i.e., they have half the rows and one column less than than Γ_N). A and B are of order $[2S + S(S-1)/2 + 1] \times [2S + S(S-1)/2 + 1]$ (i.e., they have one row and column less than $\boldsymbol{\Gamma}_N' \boldsymbol{\Gamma}_N$).

Next define $\breve{\boldsymbol{\Gamma}}_N = (\breve{\boldsymbol{\Gamma}}_N^{0\prime}, \breve{\boldsymbol{\Gamma}}_N^{1\prime})'$, which differs from $\boldsymbol{\Gamma}_N$ only by the ordering of the rows. $\breve{\boldsymbol{\Gamma}}_N^0$ corresponds to $\boldsymbol{\Gamma}_N^0$ with a zero column appended as last column, i.e., $\breve{\boldsymbol{\Gamma}}_N^0 = (\boldsymbol{\Gamma}_N^0, 0)$, such that

$$\breve{\mathbf{\Gamma}}^{0\prime}_N \breve{\mathbf{\Gamma}}^0_N = \begin{bmatrix} \breve{\mathbf{\Gamma}}^{0\prime}_N \breve{\mathbf{\Gamma}}^0_N & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a_{1,1} & & a_{1,2S+S(S-1)/2+1} & 0 \\ & \cdot & & 0 \\ a_{2S+S(S-1)/2+1,1} & & a_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \qquad \text{(B.22a)}$$

($\breve{\mathbf{\Gamma}}^{0\prime}_N \breve{\mathbf{\Gamma}}^0_N$ is of the same dimension as $\mathbf{\Gamma}'_N \mathbf{\Gamma}_N$, i.e., $[2S + S(S-1)/2 + 2] \times [2S + S(S-1)/2 + 2]$.)

$\breve{\mathbf{\Gamma}}^1_N$ is a modified version of $\mathbf{\Gamma}^1_N$, with a zero column included as second last column, such that

$$\breve{\mathbf{\Gamma}}^{1\prime}_N \breve{\mathbf{\Gamma}}^1_N = \begin{bmatrix} b_{1,1} & & 0 & b_{1,2S+S(S-1)/2+1} \\ & \cdot & 0 & \cdot \\ 0 & 0 & 0 & 0 \\ b_{2S+S(S-1)/2+1,1} & & 0 & b_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} \end{bmatrix}. \qquad \text{(B.22b)}$$

($\breve{\mathbf{\Gamma}}^{1\prime}_N \breve{\mathbf{\Gamma}}^1_N$ is of the same dimension as $\mathbf{\Gamma}'_N \mathbf{\Gamma}_N$, i.e., $[2S + S(S-1)/2 + 2] \times [2S + S(S-1)/2 + 2]$.)

Since $\breve{\mathbf{\Gamma}}_N = (\breve{\mathbf{\Gamma}}^{0\prime}_N, \breve{\mathbf{\Gamma}}^{1\prime}_N)'$ differs from $\mathbf{\Gamma}_N$ only by the ordering of the rows, it follows that

$$\mathbf{\Gamma}'_N \mathbf{\Gamma}_N = \breve{\mathbf{\Gamma}}'_N \breve{\mathbf{\Gamma}}_N = \begin{bmatrix} \breve{\mathbf{\Gamma}}^{0\prime}_N & \breve{\mathbf{\Gamma}}^{1\prime}_N \end{bmatrix} \begin{bmatrix} \breve{\mathbf{\Gamma}}^0_N \\ \breve{\mathbf{\Gamma}}^1_N \end{bmatrix} = \breve{\mathbf{\Gamma}}^{0\prime}_N \breve{\mathbf{\Gamma}}^0_N + \breve{\mathbf{\Gamma}}^{1\prime}_N \breve{\mathbf{\Gamma}}^1_N, \text{ i.e.,} \qquad \text{(B.23)}$$

$$\mathbf{\Gamma}'_N \mathbf{\Gamma}_N = \begin{bmatrix} a_{1,1} & & a_{1,2S+S(S-1)/2+1} & 0 \\ & \cdot & & 0 \\ a_{2S+S(S-1)/2+1,1} & & a_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} b_{1,1} & & 0 & b_{1,2S+S(S-1)/2+1} \\ & \cdot & 0 & \cdot \\ 0 & 0 & 0 & 0 \\ b_{2S+S(S-1)/2+1,1} & & 0 & b_{2S+S(S-1)/2+1,2S+S(S-1)/2+1} \end{bmatrix}.$$

We can thus write

$$\mathbf{x}'\mathbf{\Gamma}'_N \mathbf{\Gamma}_N \mathbf{x} = \mathbf{x}'\breve{\mathbf{\Gamma}}^{0\prime}_N \breve{\mathbf{\Gamma}}^0_N \mathbf{x} + \mathbf{x}'\breve{\mathbf{\Gamma}}^{1\prime}_N \breve{\mathbf{\Gamma}}^1_N \mathbf{x} = \mathbf{x}'_A \mathbf{A}_N \mathbf{x}_A + \mathbf{x}'_B \mathbf{B}_N \mathbf{x}_B. \qquad \text{(B.24)}$$

The vector $\mathbf{x}$ is of dimension $[2S + S(S-1)/2 + 2] \times 1$ (corresponding to the number of columns of $\mathbf{\Gamma}_N$), wheras $\mathbf{x}_A$ and $\mathbf{x}_B$ are of dimension $[2S + S(S-1)/2 + 1]$, i.e. both have one row less: $\mathbf{x}_A$ excludes the last element of $\mathbf{x}$, i.e., $\mathbf{x}_{2S+S(S-1)/2+2}$, $\mathbf{x}_B$ excludes the second-last element of $\mathbf{x}$, i.e., $\mathbf{x}_{2S+S(S-1)/2+1}$.

Again, we invoke Rao (1973, p. 62) for each quadratic form. It follows

$$\mathbf{x}_A' \mathbf{A}_N \mathbf{x}_A + \mathbf{x}_B' \mathbf{B}_N \mathbf{x}_B \geq \lambda_{\min}(\mathbf{A}_N) \mathbf{x}_A' \mathbf{x}_A + \lambda_{\min}(\mathbf{B}_N) \mathbf{x}_B' \mathbf{x}_B \geq \lambda^* (\mathbf{x}_A' \mathbf{x}_A + \mathbf{x}_B' \mathbf{x}_B) \geq \lambda^* \mathbf{x}' \mathbf{x} \qquad \text{(B.25)}$$

for any $\mathbf{x} = (x_1, x_2, ..., x_{2S+2})$.

Hence, we have shown that $\mathbf{x}' \mathbf{\Gamma}_N' \mathbf{\Gamma}_N \mathbf{x} \geq \lambda^* \mathbf{x}' \mathbf{x}$, or, equivalently,

$$\frac{\mathbf{x}' \mathbf{\Gamma}_N' \mathbf{\Gamma}_N \mathbf{x}}{\mathbf{x}' \mathbf{x}} \geq \lambda^* \quad \text{for } \mathbf{x} \neq \mathbf{0}. \qquad \text{(B.26)}$$

Next, note that in light of Rao (1973, p. 62),

$$\lambda_{\min}(\mathbf{\Gamma}_N' \mathbf{\Gamma}_N) = \inf_x \frac{\mathbf{x}' \mathbf{\Gamma}_N' \mathbf{\Gamma}_N \mathbf{x}}{\mathbf{x}' \mathbf{x}} \geq \lambda_* > 0. \qquad \text{(B.27)}$$

Using Mittelhammer (1996, p. 254) we have

$$\lambda_{\min}(\mathbf{\Gamma}_N' \mathbf{\Theta}_N \mathbf{\Gamma}_N) = \inf_x \frac{\mathbf{x}' \mathbf{\Gamma}_N' \mathbf{\Theta}_N \mathbf{\Gamma}_N x}{\mathbf{x}' \mathbf{x}} \geq \lambda_{\min}(\Xi_N^{-1}) \inf_x \frac{\mathbf{x}' \mathbf{\Gamma}_N' \mathbf{\Gamma}_N \mathbf{x}}{\mathbf{x}' \mathbf{x}}$$

$$= \lambda_{\min}(\mathbf{\Theta}_N) \lambda_{\min}(\mathbf{\Gamma}_N' \mathbf{\Gamma}_N) \geq \lambda_0 > 0, \qquad \text{(B.28)}$$

with $\lambda_0 = \bar{\lambda}_* \lambda_*$ since $= \lambda_{\min}(\mathbf{\Theta}_N) \geq \bar{\lambda}_* > 0$ by assumption (see Theorem 2).

This ensures that the true parameter vector $\mathbf{\theta}_N = (\rho_{1,N}, ..., \rho_{S,N}, \sigma_v^2, \sigma_1^2)'$ is identifiable unique.

Next note that in light of the assumptions in Theorem 2, $\mathbf{\Theta}_N$ is O(1) by the equivalence of matrix norms.

Analogous to the prove of Theorem 1, observe that $\bar{R}_N(\mathbf{\theta}) = 0$, i.e., $\bar{R}_N(\underline{\mathbf{\theta}}) = 0$ at the true parameter vector $\mathbf{\theta}_N = (\rho_{1,N}, ..., \rho_{S,N}, \sigma_v^2, \sigma_1^2)'$. It follows that

$$\overline{R}_N(\underline{\boldsymbol{\theta}}) - \overline{R}_N(\boldsymbol{\theta}_N) = (\underline{\boldsymbol{\theta}} - \boldsymbol{\theta}_N)'\boldsymbol{\Gamma}_N'\boldsymbol{\Theta}_N\boldsymbol{\Gamma}_N(\underline{\boldsymbol{\theta}} - \boldsymbol{\theta}_N). \tag{B.29}$$

Moreover, let $\mathbf{F}_N = (\tilde{\boldsymbol{\gamma}}_N, -\tilde{\boldsymbol{\Gamma}}_N)$ and $\boldsymbol{\Phi}_N = (\boldsymbol{\gamma}_N, -\boldsymbol{\Gamma}_N)$, then,

$$R_N(\underline{\boldsymbol{\theta}}) = (1, \underline{\boldsymbol{\theta}}')\mathbf{F}_N'\boldsymbol{\Theta}_N\mathbf{F}_N(1, \underline{\boldsymbol{\theta}}')' \text{ and} \tag{B.30a}$$

$$\overline{R}_N(\underline{\boldsymbol{\theta}}) = (1, \underline{\boldsymbol{\theta}}')\boldsymbol{\Phi}_N'\boldsymbol{\Theta}_N\boldsymbol{\Phi}_N(1, \underline{\boldsymbol{\theta}}')'. \tag{B.30b}$$

The remainder of the proof is now analogous to that of Theorem 1a.

Proof of Theorem 2. Asymptotic Normality of $\tilde{\boldsymbol{\theta}}_N$

To derive the asymptotic distribution of the vector $\mathbf{q}_N$, defined in (30) we invoke the central limit theorem for vectors of linear quadratic forms given by Kelejian and Prucha (2008, Theorem A.1), which is an extension of the central limit theorem for a single linear quadratic form by Kelejian and Prucha (2001, Theorem 1). The vector of quadratic forms in the present context, to which the Theorem is applied is $\mathbf{q}_N^*$. The variance-covariance matrix of $\mathbf{q}_N$ was derived above and is denoted as $\boldsymbol{\Psi}_N$. Accordingly, the variance-covariance matrix of $\mathbf{q}_N^* = N^{1/2}\mathbf{q}_N$ is given by $\boldsymbol{\Psi}_N^* = N\boldsymbol{\Psi}_N$ and $(\boldsymbol{\Psi}_N^*)^{-1/2} = N^{-1/2}\boldsymbol{\Psi}_N^{-1/2}$.

Note that in light of Assumptions 1, 2 and 7 (and Lemma B.1), the stacked innovations $\boldsymbol{\xi}_N$, the matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s = 1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, and the vectors $\mathbf{a}_{1,s,N}$, …, $\mathbf{a}_{4,s,N}$, $s = 1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$ satisfy the assumptions of central limit theorem by Kelejian and Prucha (2008, Theorem A.1). In the application of Theorem A.1, note that the sample size is given by $NT + N = N(T+1)$ rather than N. As Kelejian and Prucha (2001, p. 227, fn. 13) point out, Theorem A.1 "also holds if the sample size is taken to be k_n rather than n (with $k_n \uparrow \infty$ as $N \to \infty$)." In the present case we have $K_N = (T+1)N$, with $T \geq 1$ and fixed, which ensures that $K_n \uparrow \infty$ as $N \to \infty$. Consequently, Theorem A.1 still applies to each quadratic form in $\mathbf{q}_N^*$. Moreover, as can be observed from the proof of Theorem A.1 in Kelejian and Prucha (2008), the extension of the Theorem from a scalar to a vector of vector of quadratic forms holds up under by this alternative definition of the sample size.

It follows that

$$-(\boldsymbol{\Psi}_N^*)^{-1/2}\mathbf{q}_N^* = -N^{-1/2}\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N^* = -\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} (\mathbf{0}, \mathbf{I}_{4S+2}), \tag{B.31}$$

since $N^{-1}\lambda_{\min}(\boldsymbol{\Psi}_N^*) = N^{-1}\lambda_{\min}(N\boldsymbol{\Psi}_N) = \lambda_{\min}(\boldsymbol{\Psi}_N) > 0$ by assumption as required in Theorem A.1.

Since the row and column sums of the matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s = 1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, and the vectors $\mathbf{a}_{1,s,N}$, …, $\mathbf{a}_{4,s,N}$, $s = 1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$, and the variances $\sigma_{v,N}^2$ and $\sigma_{\mu,N}^2$ are bounded uniformly in absolute value, it follows in light of (38) that the elements of $\boldsymbol{\Psi}_N$ and also those of $\boldsymbol{\Psi}_N^{1/2}$ are bounded uniformly in absolute value.

We next turn to the derivation of the limiting distribution of the GM estimator $\tilde{\boldsymbol{\theta}}_N$. In Theorem 1 we showed that the GM estimator $\tilde{\boldsymbol{\theta}}_N$ defined by (18) is consistent. It follows that – apart from a set of the sample space whose probability tends to zero – the estimator satisfies the following first order condition:[6]

$$\frac{\partial}{\partial \boldsymbol{\rho}}\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)'\tilde{\boldsymbol{\Theta}}_N\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N) = \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'}\tilde{\boldsymbol{\Theta}}_N\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N) = \mathbf{0}, \tag{B.32}$$

which is a $(S+2)\times 1$ vector, the rows corresponding the partial derivatives of the criterion function with respect to $\rho_{s,N}$, $s = 1,...,S$, σ_v^2, and σ_1^2.

Substituting the mean value theorem expression

$$\mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N) = \mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N) + \frac{\partial \mathbf{q}_N(\bar{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N), \tag{B.33}$$

where $\bar{\boldsymbol{\theta}}_N$ is some between value, into the first-order condition yields

$$\frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\Theta}}_N\frac{\partial \mathbf{q}_N(\bar{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'}N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = -\frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\Theta}}_N N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N,\boldsymbol{\Delta}_N). \tag{B.34}$$

Observe that $\frac{\partial \mathbf{q}_N(\boldsymbol{\theta},\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'} = \tilde{\boldsymbol{\Gamma}}_N\mathfrak{B}_N$ and consider the two $(S+2)\times(S+2)$ matrices

$$\tilde{\boldsymbol{\Xi}}_N = \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}}\tilde{\boldsymbol{\Theta}}_N\frac{\partial \mathbf{q}_N(\bar{\boldsymbol{\theta}}_N,\boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}'} = \tilde{\mathfrak{B}}_N'\tilde{\boldsymbol{\Gamma}}_N'\tilde{\boldsymbol{\Theta}}_N\tilde{\boldsymbol{\Gamma}}_N\overline{\mathfrak{B}}_N, \tag{B.35}$$

[6] The leading two and the negative sign are ignored without further consequences for the proof.

$$\boldsymbol{\Xi}_N = \boldsymbol{\mathfrak{B}}_N' \boldsymbol{\Gamma}_N' \boldsymbol{\Theta}_N \boldsymbol{\Gamma}_N \boldsymbol{\mathfrak{B}}_N , \tag{B.36}$$

where $\tilde{\boldsymbol{\mathfrak{B}}}_N$ and $\overline{\boldsymbol{\mathfrak{B}}}_N$ correspond to $\boldsymbol{\mathfrak{B}}_N$ as defined above with $\tilde{\boldsymbol{\theta}}_N$ and $\bar{\boldsymbol{\theta}}_N$ substituted for $\boldsymbol{\theta}_N$. Notice that $\boldsymbol{\Xi}_N$ is positive definite, since $\boldsymbol{\Gamma}_N$ and $\boldsymbol{\Theta}_N$ are positive definite by assumption and the $[2S + S(S-1)/2 + 2] \times (S+2)$ matrix $\boldsymbol{\mathfrak{B}}_N$ has full column rank.

In the proof of Theorem 1 (and Lemma B.1) we have demonstrated that $\tilde{\boldsymbol{\Gamma}}_N - \boldsymbol{\Gamma}_N \xrightarrow{p} \mathbf{0}$ and that the elements of $\boldsymbol{\Gamma}_N$ and $\tilde{\boldsymbol{\Gamma}}_N$ are $O(1)$ and $O_p(1)$, respectively. By Assumption 5, $\tilde{\boldsymbol{\Theta}}_N - \boldsymbol{\Theta}_N = o_p(1)$, $\boldsymbol{\Theta}_N = O(1)$ and $\tilde{\boldsymbol{\Theta}}_N = O_p(1)$. Since $\tilde{\boldsymbol{\rho}}_N$ and $\bar{\boldsymbol{\rho}}_N$ (and thus also $\tilde{\boldsymbol{\mathfrak{B}}}_N$ and $\overline{\boldsymbol{\mathfrak{B}}}_N$) are consistent and bounded uniformly in probability, if follows that $\tilde{\boldsymbol{\Xi}}_N - \boldsymbol{\Xi}_N = o_p(1)$, $\tilde{\boldsymbol{\Xi}}_N = O_p(1)$, and $\boldsymbol{\Xi}_N = O(1)$. Moreover, $\boldsymbol{\Xi}_N$ is positive definite and thus invertible, and its inverse $\boldsymbol{\Xi}_N^{-1}$ is also $O(1)$.

Denote $\tilde{\boldsymbol{\Xi}}_N^{+}$ as the generalized inverse of $\tilde{\boldsymbol{\Xi}}_N$. It then follows as a special case of Lemma F1 in Pötscher and Prucha (1997) that $\tilde{\boldsymbol{\Xi}}_N$ is non-singular with probability approaching 1 as $N \to \infty$, that $\tilde{\boldsymbol{\Xi}}_N^{+}$ is $O_p(1)$, and that $\tilde{\boldsymbol{\Xi}}_N^{+} - \boldsymbol{\Xi}_N^{-1} = o_p(1)$.

Pre-multiplying (B.34) with $\tilde{\boldsymbol{\Xi}}_N^{+}$ we obtain, after rearranging terms,

$$N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = (\mathbf{I}_{S+2} - \tilde{\boldsymbol{\Xi}}_N^{+}\tilde{\boldsymbol{\Xi}}_N) N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) - N^{1/2}\tilde{\boldsymbol{\Xi}}_N^{+} \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}} \tilde{\boldsymbol{\Theta}}_N \mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N) . \tag{B.37}$$

In light of the discussion above, the first term on the right-hand side is zero on ω-sets of probability approaching 1 (compare Pötscher and Prucha, 1997, pp. 228). This yields

$$N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = -\tilde{\boldsymbol{\Xi}}_N^{+} \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}} \tilde{\boldsymbol{\Theta}}_N N^{1/2} \mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N) + o_p(1) . \tag{B.38}$$

Next observe that

$$\tilde{\boldsymbol{\Xi}}_N^{+} \frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}} \tilde{\boldsymbol{\Theta}}_N - \boldsymbol{\Xi}_N^{-1} \boldsymbol{\mathfrak{B}}_N' \boldsymbol{\Gamma}_N' \boldsymbol{\Theta}_N = o_p(1) , \tag{B.39}$$

since $\tilde{\boldsymbol{\Xi}}_N^{+} - \boldsymbol{\Xi}_N^{-1} = o_p(1)$ and $\frac{\partial \mathbf{q}_N(\tilde{\boldsymbol{\theta}}_N, \boldsymbol{\Delta}_N)}{\partial \boldsymbol{\theta}} - \boldsymbol{\mathfrak{B}}_N' \boldsymbol{\Gamma}_N' = o_p(1)$.

As we showed in section III, the elements of $N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N)$ can be expressed as

$$N^{1/2}\mathbf{q}_N(\boldsymbol{\theta}_N, \boldsymbol{\Delta}_N) = N^{-1/2}\mathbf{q}_N^* + o_p(1) = \mathbf{q}_N + o_p(1). \tag{B.40}$$

where $\mathbf{q}_N^*$ is defined in (27), and that

$$-(\boldsymbol{\Psi}_N^*)^{-1/2}\mathbf{q}_N^* = -N^{-1/2}\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N^* = -\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} (\mathbf{0}, \mathbf{I}_{4S+2}). \tag{B.41}$$

It now follows from (B.38), (B.39), and (B.40) that

$$N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = \boldsymbol{\Xi}_N^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N^{1/2}(-\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N) + o_p(1). \tag{B.42}$$

Since all nonstochastic terms on the right hand side from (B.42) are $O(1)$ it follows that $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$ is $O_p(1)$. To derive the asymptotic distribution of $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$, we invoke Corollary F4 in Pötscher and Prucha (1997). In the present context, we have

$$\boldsymbol{\zeta}_N = -\boldsymbol{\Psi}_N^{-1/2}\mathbf{q}_N \xrightarrow{d} \boldsymbol{\zeta} \sim N(\mathbf{0}, \mathbf{I}_{4S+2}),$$
$$N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = \boldsymbol{\mathfrak{A}}_N\boldsymbol{\zeta}_N + o_p(1) \text{ , with}$$
$$\boldsymbol{\mathfrak{A}}_N = \boldsymbol{\Xi}_N^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N^{1/2}.$$

Furthermore, $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N) = O_p(1)$ and its variance-covariance matrix is

$$\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\boldsymbol{\Theta}_N) = (\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N\boldsymbol{\Theta}_N\mathbf{J}_N(\mathbf{J}'_N\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1},$$

where $\boldsymbol{\Omega}_{\tilde{\boldsymbol{\theta}}_N}$ is positive definite.

As a final point it has to be shown that $\liminf_{N\to\infty}\lambda_{\min}(\boldsymbol{\mathfrak{A}}_N\boldsymbol{\mathfrak{A}}'_N) > 0$ as required in Corollary F4 in Pötscher and Prucha (1997). Observe that

$$\begin{aligned}\lambda_{\min}(\boldsymbol{\mathfrak{A}}_N\boldsymbol{\mathfrak{A}}'_N) &= \lambda_{\min}(\boldsymbol{\Xi}_N^{-1}\mathbf{J}'_N\boldsymbol{\Theta}_N\boldsymbol{\Psi}_N\boldsymbol{\Theta}'_N\mathbf{J}_N\boldsymbol{\Xi}_N^{-1}) \\ &\geq \lambda_{\min}(\boldsymbol{\Psi}_N)\lambda_{\min}(\boldsymbol{\Theta}_N\boldsymbol{\Theta}'_N)\lambda_{\min}(\boldsymbol{\Xi}_N^{-1}\boldsymbol{\Xi}_N^{-1})\lambda_{\min}(\boldsymbol{\Gamma}'_N\boldsymbol{\Gamma}_N)\lambda_{\min}(\boldsymbol{\mathfrak{B}}'_N\boldsymbol{\mathfrak{B}}_N) > 0,\end{aligned} \tag{B.43}$$

since the matrices involved are all positive definite.

Consistency Proof for Estimates of Third and Fourth Moments of the Error Components

Consistent estimates for the second moments of $v_{it,N}$ and $\mu_{i,N}$ are delivered by the GM estimators defined in (17) and (18), respectively (See Theorems 1a and 1b). In the following we prove that the estimators for the third and fourth moments of $v_{it,N}$ and $\mu_{it,N}$, defined in (39) and (40) are also consistent. The proof draws on Gilbert (2002), who considers the estimation of third and fourth moments in error component models without spatial lags and without spatial regressive disturbances. For reasons that will become clear below, we depart from the convention adopted so far to use indexation $i = 1,...,NT$ for the stacked series. In the subsequent proof we use the double indexation it, with $i = 1,...,N$, $t = 1,...,T$.

Preliminary Remarks

Note first that

$$\tilde{\boldsymbol{\varepsilon}}_N = \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N \text{, where} \tag{B.44}$$

$$\boldsymbol{\eta}_N = [\sum_{m=1}^{S}(\rho_{m,N} - \tilde{\rho}_{m,N})\mathbf{I}_T \otimes \mathbf{M}_{m,N}][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N$$
$$+ [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{D}_N\boldsymbol{\Delta}_N + [\sum_{m=1}^{S}(\rho_{m,N} - \tilde{\rho}_{m,N})(\mathbf{I}_T \otimes \mathbf{M}_{m,N})]\mathbf{D}_N\boldsymbol{\Delta}_N .$$

This can also be written as

$$\boldsymbol{\eta}_N = \mathbf{R}_N\mathfrak{g}_N , \tag{B.45}$$

where

$$\mathbf{R}_N = (\mathbf{R}_{1,N}, \mathbf{R}_{2,N}, \mathbf{R}_{3,N}) \text{ with}$$

$$\mathbf{R}_{1,N} = \mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})\mathbf{D}_N ,$$

$$\mathbf{R}_{2,N} = \{(\mathbf{I}_T \otimes \mathbf{M}_{1,N})[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N, ..., (\mathbf{I}_T \otimes \mathbf{M}_{S,N})[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N\},$$

$$\mathbf{R}_{3,N} = [(\mathbf{I}_T \otimes \mathbf{M}_{1,N})\mathbf{D}_N, ..., (\mathbf{I}_T \otimes \mathbf{M}_{S,N})\mathbf{D}_N] ,$$

$$\mathfrak{g}_N = \begin{bmatrix} \boldsymbol{\Delta}_N \\ (\boldsymbol{\rho}_N - \tilde{\boldsymbol{\rho}}_N) \\ (\boldsymbol{\rho}_N - \tilde{\boldsymbol{\rho}}_N) \otimes \boldsymbol{\Delta}_N \end{bmatrix}.$$

In light of Assumption 3 and since the elements of $\mathbf{D}_N = (\mathbf{d}'_{1.,N},...,\mathbf{d}'_{N.,N})'$ have bounded fourth moments (by Assumption in Theorem 3), each column of the matrix $\mathbf{R}_N$ is of the form $\underline{\boldsymbol{\pi}}_N + \underline{\boldsymbol{\Pi}}_N \boldsymbol{\xi}_N$, where the elements of the $NT \times 1$ vector $\underline{\boldsymbol{\pi}}_N$ are bounded uniformly in absolute value by some finite constant, the row and column sums of the $NT \times NT$ matrix $\underline{\boldsymbol{\Pi}}_N$ are bounded uniformly in absolute value by some finite constant, and the fourth moments of the elements of $\boldsymbol{\xi}_N$ are also bounded by some finite constant. It follows that the fourth moments of the elements of $\mathbf{R}_N$ are also bounded by some finite constant by Lemma C.2 in Kelejian and Prucha (2008).[7]

As a consequence, $\|\boldsymbol{\eta}_N\| \leq \|\mathbf{R}_N\| \|\mathfrak{g}_N\|$, or for the *i*-th element of the $NT \times 1$ vector $\boldsymbol{\eta}_N$,

$$|\eta_{i,N}| \leq \|\mathfrak{g}_N\| \|\mathbf{r}_{i.,N}\| = \alpha_N \beta_{i,N}, \tag{B.46}$$

where $\mathbf{r}_{i.,N}$ denotes the *i*-th row of $\mathbf{R}_N$, $\alpha_N = \|\mathfrak{g}_N\|$, and $\beta_{i,N} = \|\mathbf{r}_{i.,N}\|$ with $E|\beta_{i,N}|^4 \leq K_\beta < \infty$. Without loss of generality we can select K_β such that $E(\beta_{i,N}^\gamma) \leq K_\beta$ for $\gamma \leq 4$. By Assumption 1 there is also some K_ε such that $E|\varepsilon_{i,N}|^\gamma \leq K_\varepsilon < \infty$ for $\gamma \leq 4$. In the following we use K to denote the larger bound, i.e., $K = \max(K_\varepsilon, K_\beta)$. Also note that $N^{1/2}\alpha_N = O_p(1)$.

Estimation of Third Moments

We first consider $\sigma_\varepsilon^{(3)} = E(\varepsilon_{it,N}^3)$ and its estimate $\tilde{\sigma}_{\varepsilon,N}^{(3)} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\varepsilon}_{it,N}^3$.

Using (B.44) we have

$$\begin{aligned}
\tilde{\sigma}_{\varepsilon,N}^{(3)} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N} + \eta_{it,N})^3 \qquad \text{(B.47)}\\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N}^3 + 3\varepsilon_{it,N}^2\eta_{it,N} + 3\eta_{it,N}^2\varepsilon_{it,N} + \eta_{it,N}^3)\\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it,N}^3 + \frac{3}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it,N}^2\eta_{it,N} + \frac{3}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\eta_{it,N}^2\varepsilon_{it,N} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\eta_{it,N}^3\\
&= \varphi_{1,N} + \varphi_{2,N} + \varphi_{3,N} + \varphi_{4,N}.
\end{aligned}$$

By the weak law of large numbers for i.i.d. random variables, $\varphi_{1,N}$ converges to $\sigma_\varepsilon^{(3)}$ as $N \to \infty$.

[7] See also Remark A.1 in Appendix A.

Next consider

$$\varphi_{2,N} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|3\varepsilon_{it,N}^2\eta_{it,N}\right| \tag{B.48}$$
$$\leq \frac{3}{NT}\alpha_N\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\varepsilon_{it,N}^2\beta_{it,N}\right|$$
$$\leq \frac{3}{NT}\alpha_N\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\varepsilon_{it,N}\right|^4\right)^{1/2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\beta_{it,N}\right|^2\right)^{1/2}$$
$$= 3(N^{1/2}\alpha_N)N^{-1/2}\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\varepsilon_{it,N}\right|^4\right)^{1/2}\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\beta_{it,N}\right|^2\right)^{1/2} = o_p(1),$$

since $E\left|\varepsilon_{it,N}\right|^4 \leq K < \infty$ and thus $\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\varepsilon_{it,N}\right|^4\right)^{1/2} = O_p(1)$, $E\left|\beta_{it,N}\right|^2 \leq K < \infty$ and thus $\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\beta_{it,N}\right|^2\right)^{1/2} = O_p(1)$, $N^{1/2}\alpha_N = O_p(1)$, and $N^{-1/2} = o(1)$.

Observe further that

$$\varphi_{3,N} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|3\eta_{it,N}^2\varepsilon_{it,N}\right| \tag{B.49}$$
$$\leq \frac{3}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\alpha_N^2\beta_{i,N}^2\varepsilon_{it,N}\right| = \frac{3}{NT}\alpha_N^2\sum_{i=1}^{NT}\left|\beta_{i,N}^2\varepsilon_{it,N}\right|$$
$$= 3(N^{1/2}\alpha_N)^2N^{-1}\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\varepsilon_{it,N}\right|^2\right)^{1/2}\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\beta_{i,N}\right|^4\right)^{1/2} = o_p(1),$$

since $E\left|\varepsilon_{it,N}\right|^2 \leq K < \infty$ and thus $\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\varepsilon_{it,N}\right|^2\right)^{1/2} = O_p(1)$, $E\left|\beta_{i,N}\right|^4 \leq K < \infty$ and thus $\left((NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\beta_{i,N}\right|^4\right)^{1/2} = O_p(1)$, $N^{1/2}\alpha_N = O_p(1)$, and $N^{-1} = o(1)$.

Finally,

$$\varphi_{4,N} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\eta_{it,N}^3\right| \tag{B.50}$$
$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\alpha_N\right|^3\left|\beta_{i,N}\right|^3$$

$$\leq \frac{1}{NT}(N^{1/2}\alpha_N)^3 N^{-3/2} N\left(N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\beta_{i,N}\right|^3 \right)$$

$$= (N^{1/2}\alpha_N)^3 N^{-3/2}\left((NT)^{-1}\sum_{i=1}^{NT}\left|\beta_{i,N}\right|^3 \right) = o_p(1),$$

since $E(\left|\beta_{i,N}\right|^3) \leq K < \infty$ and thus $\left((NT)^{-1}\sum_{i=1}^{NT}\left|\beta_{i,N}\right|^3 \right) = O_p(1)$, $N^{1/2}\alpha_N = O_p(1)$, and $N^{-3/2} = o(1)$. As a consequence, we have $\tilde{\sigma}^{(3)}_{\varepsilon,N} - \sigma^{(3)}_{\varepsilon} = o_p(1)$, $\sigma^{(3)}_{\varepsilon} = O(1)$ by Assumption 1, and $\tilde{\sigma}^{(3)}_{\varepsilon,N} = O_p(1)$.

Next consider the third moments of the unit-specific error component $\sigma^{(3)}_{\mu}$ and its estimate $\tilde{\sigma}^{(3)}_{\mu}$, which can be expressed (compare Gilbert, 2002, p. 48ff.) as

$$\sigma^{(3)}_{\mu} = E(\varepsilon_{is,N}\varepsilon^2_{it,N}) \text{ for any given } i \text{ and } s \neq t, \tag{B.50a}$$

$$\tilde{\sigma}^{(3)}_{\mu,N} = \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}\tilde{\varepsilon}_{is,N}\tilde{\varepsilon}^2_{it,N}. \tag{B.50b}$$

Notice that by Assumption 1, $\sigma^{(3)}_{\mu}$ is invariant to the choice of *i, s* and *t*. Using (B.44), we have

$$\tilde{\sigma}^{(3)}_{\mu,N} = \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}(\varepsilon_{is,N}+\eta_{is,N})(\varepsilon_{it,N}+\eta_{it,N})^2 \tag{B.51}$$

$$= \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}(\varepsilon_{is,N}\varepsilon^2_{it,N} + 2\eta_{it,N}\varepsilon_{is,N} + \eta^2_{it,N}\varepsilon_{is,N} + \eta_{is,N}\varepsilon^2_{it,N} + 2\eta_{is,N}\eta_{it,N} + \eta_{is,N}\eta^2_{it,N})$$

$$= \phi_{1,N} + \phi_{2,N} + \phi_{3,N} + \phi_{4,N} + \phi_{5,N} + \phi_{6,N}$$

Observe that

$$\phi_{1,N} = \sum_{i=1}^{N}\sum_{s}^{T}\sum_{t\neq s}^{T}\varepsilon_{is,N}\varepsilon^2_{it,N} = \sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}(\mu_{is,N}+v_{is,N})(\mu_{it,N}+v_{it,N})^2 \tag{B.52}$$

$$= \sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}(\mu^3_{i,N} + 2\mu^2_{i,N}v_{it,N} + \mu_{i,N}v^2_{it,N} + v_{is,N}\mu^2_{i,N} + 2\mu_{i,N}v_{it,N}v_{is,N} + v^2_{it,N}v_{is,N})$$

$$= \phi_{11,N} + \phi_{12,N} + \phi_{13,N} + \phi_{14,N} + \phi_{15,N} + \phi_{16,N}.$$

By the weak law of large numbers $\phi_{11,N}$ converges in probability to $\sigma_\mu^{(3)}$. Notice further that, by the properties of $v_{it,N}$ and $\mu_{it,N}$ (see Assumption 1), $\phi_{12,N}, \phi_{13,N}, \phi_{14,N}, \phi_{15,N}$, and $\phi_{16,N}$ are all $o_p(1)$. As a consequence, $\phi_{1,N}$ converges in probability to $\sigma_\mu^{(3)}$.

Next observe that

$$\phi_{2,N} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} 2\eta_{it,N}\varepsilon_{is,N} \tag{B.53}$$

$$\leq \frac{2}{NT(T-1)} \alpha_N \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \varepsilon_{is,N}\beta_{it,N}$$

$$\leq 2N^{1/2}\alpha_N N^{-1/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \left|\varepsilon_{is,N}\right|^2 \right)^{1/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{it,N}^2 \right)^{1/2} = o_p(1),$$

$$\phi_{3,N} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \eta_{it,N}^2 \varepsilon_{is,N} \tag{B.54}$$

$$\leq (N^{1/2}\alpha_N)^2 N^{-1} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \left|\varepsilon_{is,N}\right|^2 \right)^{1/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{it,N}^4 \right)^{1/2} = o_p(1),$$

$$\phi_{4,N} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \eta_{is,N}\varepsilon_{it,N}^2 \tag{B.55}$$

$$\leq N^{1/2}\alpha_N N^{-1/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s}^{T} \sum_{t \neq s}^{T} \left|\varepsilon_{is,N}\right|^4 \right)^{1/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s}^{T} \sum_{t \neq s}^{T} \beta_{it,N}^2 \right)^{1/2} = o_p(1),$$

$$\phi_{5,N} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s}^{T} \sum_{t \neq s}^{T} 2\eta_{is,N}\eta_{it,N} \tag{B.56}$$

$$\leq \frac{2}{NT(T-1)} (N^{1/2}\alpha_N)^2 N^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{is,N}\beta_{it,N}$$

$$\leq 2(N^{1/2}\alpha_N)^2 N^{-1} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{is,N}^2 \right) \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{it,N}^2 \right)^{1/2} = o_p(1),$$

$$\phi_{6,N} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \eta_{is,N} \eta_{it,N}^2 \tag{B.57}$$

$$\leq 2(N^{1/2}\alpha_N)^3 N^{-3/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{is,N}^2 \right)^{1/2} \left([NT(T-1)]^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \beta_{it,N}^4 \right)^{1/2} = o_p(1),$$

because $N^{1/2}\alpha_N$ is $O_p(1)$ and the terms in brackets expressions are all $O_p(1)$, since $E|\varepsilon_{is,N}|^{\gamma} < K < \infty$ and $E|\beta_{it,N}|^{\gamma} < K < \infty$ for $\gamma \leq 4$ and all *N*. It follows that $\tilde{\sigma}_{\mu,N}^{(3)} - \sigma_{\mu}^{(3)} = o_p(1)$, $\sigma_{\mu}^{(3)} = O(1)$ by Assumption 1, and that $\tilde{\sigma}_{\mu,N}^{(3)} = O_p(1)$. Obviously, we then also have that $(\tilde{\sigma}_{\varepsilon,N}^{(3)} - \tilde{\sigma}_{\mu,N}^{(3)}) - \sigma_v^{(3)} = \tilde{\sigma}_{v,N}^{(3)} - \sigma_v^{(3)} = o_p(1)$.

Estimation of Fourth Moments

Consider the fourth moment of $\mu_{i,N}$ and its estimate, which can be expressed as (compare Gilbert, 2002, p. 48ff.):

$$\sigma_{\mu}^4 = E(\varepsilon_{is}\varepsilon_{it}^3) - 3E(\varepsilon_{is}\varepsilon_{it})[E(\varepsilon_{it}^2) - E(\varepsilon_{is}\varepsilon_{it})] \text{ for any given } i \text{ and } s \neq t, \tag{B.58a}$$

$$\tilde{\sigma}_{\mu,N}^{(4)} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \tilde{\varepsilon}_{is,N} \tilde{\varepsilon}_{it,N}^3 \tag{B.58b}$$

$$- \frac{3}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \tilde{\varepsilon}_{is,N} \tilde{\varepsilon}_{it,N} \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{\varepsilon}_{it,N}^2 - \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} \tilde{\varepsilon}_{is,N} \tilde{\varepsilon}_{it,N} \right)$$

$$= \delta_{1,N} - \delta_{2,N}(\tilde{\sigma}_{\varepsilon,N}^2 - \delta_{2,N}).$$

Observe that

$$\delta_{1,N} = \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} (\varepsilon_{is,N} + \eta_{is,N})(\varepsilon_{it,N} + \eta_{it,N})^3 \tag{B.59}$$

$$= \frac{1}{NT(T-1)} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{\substack{t=1 \\ t \neq s}}^{T} (\varepsilon_{is,N}\varepsilon_{it,N}^3 + 3\varepsilon_{is,N}\varepsilon_{it,N}^2\eta_{it,N} + 3\eta_{it,N}^2\varepsilon_{is,N}\varepsilon_{it,N} + \varepsilon_{is,N}\eta_{it,N}^3$$

$$+ \eta_{is,N}\varepsilon_{it,N}^3 + 3\varepsilon_{it,N}^2\eta_{it,N}\eta_{is,N} + 3\eta_{it,N}^2\eta_{is,N}\varepsilon_{it,N} + \eta_{it,N}^3\eta_{is,N})$$

$$= \delta_{11,N} + \delta_{12,N} + \delta_{13,N} + \delta_{14,N} + \delta_{15,N} + \delta_{16,N} + \delta_{17,N} + \delta_{18,N}.$$

The first term $\delta_{11,N}$ can also be written as

$$\delta_{11,N} = \varepsilon_{is,N}\varepsilon_{it,N}^3 = (\mu_{is,N} + v_{is,N})(\mu_{it,N} + v_{it,N})^3 = \tag{B.60}$$
$$= (\mu_{is,N} + v_{is,N})(\mu_{it,N}^3 + 3\mu_{it,N}^2 v_{it,N} + 3\mu_{it,N} v_{it,N}^2 + v_{it,N}^3)$$
$$= (\mu_{is,N}\mu_{it,N}^3 + 3\mu_{is,N}\mu_{it,N}^2 v_{it,N} + 3\mu_{is,N}\mu_{it,N} v_{it,N}^2 + \mu_{is,N} v_{it,N}^3$$
$$+ v_{is,N}\mu_{it,N}^3 + 3\mu_{it,N}^2 v_{is,N} v_{it,N} + 3\mu_{it,N} v_{is,N} v_{it,N}^2 + v_{is,N} v_{it,N}^3)$$
$$= (\mu_{i,N}^4 + 3\mu_{i,N}^3 v_{it,N} + 3\mu_{i,N}^2 v_{it,N}^2 + \mu_{is,N} v_{it,N}^3 + v_{is,N}\mu_{it,N}^3 + 3\mu_{it,N}^2 v_{is,N} v_{it,N} + 3\mu_{it,N} v_{is,N} v_{it,N}^2 + v_{is,N} v_{it,N}^3).$$

By the properties of $v_{it,N}$ and $\mu_{it,N}$ (see Assumption 1), $\delta_{11,N}$ converges in probability to $\sigma_\mu^{(4)} + 3\sigma_\mu^2\sigma_v^2$.

Moreover, it follows from the properties of $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ (see Assumption 1), that the terms $\delta_{12,N}, \delta_{13,N}, \delta_{14,N}, \delta_{15,N}, \delta_{16,N}, \delta_{17,N}, \delta_{18,N}$ are $o_p(1)$. It follows that $\delta_{1,N}$ converges in probability to $\sigma_\mu^{(4)} + 3\sigma_\mu^2\sigma_v^2$.

Next consider

$$\delta_{2,N} = \frac{3}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{t\neq s}^{T}(\varepsilon_{is,N} + \eta_{it,N})(\varepsilon_{it,N} + \eta_{it,N}) \tag{B.61}$$
$$= \frac{3}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{t\neq s}^{T}(\varepsilon_{is,N} + \eta_{is,N})(\varepsilon_{it,N} + \eta_{it,N})$$
$$= \frac{3}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{t\neq s}^{T}(\varepsilon_{is,N}\varepsilon_{it,N} + \eta_{is,N}\varepsilon_{it,N} + \varepsilon_{is,N}\eta_{it,N} + \eta_{is,N}\eta_{it,N}),$$

which converges to σ_μ^2 by the weak law of large numbers, since $E[\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is,N}\frac{1}{(T-1)}\sum_{t\neq s}^{T}\varepsilon_{it,N}] = \sigma_\mu^2$ for $s \neq t$ by the properties of $v_{it,N}$ and $\mu_{it,N}$ and the sum over the remainder terms appearing in $\delta_{2,N}$ are $o_p(1)$ by arguments analogous to those for $\phi_{2,N}$ and $\phi_{5,N}$ (see B.53 and B.56). Finally, by arguments similar to the proof for the consistency of the estimate of the third moment, it follows that $\tilde{\sigma}_{\varepsilon,N}^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\varepsilon}_{it,N}^2$ converges in probability to σ_ε^2. As a consequence, $\tilde{\sigma}_{\mu,N}^{(4)} - \sigma_\mu^{(4)} = o_p(1)$, $\sigma_\mu^{(4)} = O(1)$ by Assumption 1, and that $\tilde{\sigma}_{\mu,N}^{(4)} = O_p(1)$.

We next consider the fourth moment of $v_{it,N}$ and its estimate, which can be written as (compare Gilbert, 2002, p. 48ff.):

$$\sigma_v^{(4)} = E(\varepsilon_{it}^4) - E(\varepsilon_{is}\varepsilon_{it}^3) - 3E(\varepsilon_{is}\varepsilon_{it})[E(\varepsilon_{it}^2) - E(\varepsilon_{is}\varepsilon_{it})] \text{ for any given } i \text{ and } s \neq t, \text{ (B.62a)}$$

$$\tilde{\sigma}_{v,N}^{(4)} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\tilde{\varepsilon}_{it,N}^4 - \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}\tilde{\varepsilon}_{is,N}\tilde{\varepsilon}_{it,N}^3 \qquad \text{(B.62b)}$$

$$-\frac{3}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}\tilde{\varepsilon}_{is,N}\tilde{\varepsilon}_{it,N}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\varepsilon}_{it,N}^2 - \frac{1}{NT(T-1)}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\substack{t=1 \\ t\neq s}}^{T}\tilde{\varepsilon}_{is,N}\tilde{\varepsilon}_{it,N}\right)$$

$$= \chi_{1,N} - \delta_{1,N} - \delta_{2,N}(\tilde{\sigma}_{\varepsilon,N}^2 - \delta_{2,N}).$$

We have already shown that $\delta_{1,N}$ converges in probability to $\sigma_\mu^{(4)} + 3\sigma_\mu^2\sigma_v^2$ and that $3\delta_{2,N}(\tilde{\sigma}_{\varepsilon,N}^2 - \delta_{2,N})$ converges in probability to $3\sigma_\mu^2\sigma_v^2$. Next, expanding $\chi_{1,N}$, we obtain

$$\chi_{1,N} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N} + \eta_{it,N})^4 \qquad \text{(B.63)}$$

$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N}^2 + 2\varepsilon_{it,N}\eta_{it,N} + \eta_{it,N}^2)(\varepsilon_{it,N}^2 + 2\varepsilon_{it,N}\eta_{it,N} + \eta_{it,N}^2)$$

$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N}^4 + 2\varepsilon_{it,N}^3\eta_{it,N} + \varepsilon_{it,N}^2\eta_{it,N}^2)$$

$$+ \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(2\varepsilon_{it,N}^3\eta_{it,N} + 4\varepsilon_{it,N}^2\eta_{it,N}^2 + 2\varepsilon_{it,N}\eta_{it,N}^3)$$

$$+ \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N}^2\eta_{it,N}^2 + 2\varepsilon_{it,N}\eta_{it,N}^3 + \eta_{it,N}^4)$$

$$= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it,N}^4 + 4\varepsilon_{it,N}^3\eta_{it,N} + 6\varepsilon_{it,N}^2\eta_{it,N}^2 + 4\varepsilon_{it,N}\eta_{it,N}^3 + \eta_{it,N}^4)$$

Observing that $\frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\varepsilon_{it,N}^4$ converges in probability to $\sigma_\mu^{(4)} + \sigma_v^{(4)} + 6\sigma_\mu^2\sigma_v^2$ and that the remainder terms of $\chi_{1,N}$ are all $o_p(1)$, it follows that $\tilde{\sigma}_{v,N}^{(4)} - \sigma_v^{(4)} = o_p(1)$, $\sigma_v^{(4)} = O(1)$ by Assumption 1, and that $\tilde{\sigma}_{v,N}^{(4)} = O_p(1)$.

III. Proof of Theorem 3 (Variance-Covariance Estimation)

Lemma B.2

Suppose Assumptions 1-4 hold. Furthermore, assume that $\sup_N \sum_{m=1}^{S}\left|\rho_{m,N}\right| < 1$, and that the row and column sums of $\mathbf{M}_{m,N}$, $m = 1,...,S$ are bounded uniformly in absolute value by 1 and some finite constant respectively. Let $\tilde{\sigma}_v^2$, $\tilde{\sigma}_1^2$, and $\tilde{\rho}_{s,N}$, $s = 1,...,S$, be estimators, satisfying

$\tilde{\sigma}_v^2 - \sigma_v^2 = o_p(1)$, $\tilde{\sigma}_1^2 - \sigma_1^2 = o_p(1)$, $\tilde{\rho}_{s,N} - \rho_{s,N} = o_p(1)$, $s = 1,...,S$. Let the $NT \times NT$ or $N \times N$ matrix $\mathbf{F}_N$ be of the form (compare Lemmata 1 and 2):

(a) $\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$,

$\mathbf{F}_{\mu,N} = (\mathbf{e}'_T \otimes \mathbf{I}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$,

(b) $\mathbf{F}^{**}_{v,N} = (\sigma_v^{-2}\mathbf{Q}_{0,N} + \sigma_1^{-2}\mathbf{Q}_{1,N})\mathbf{H}^*_N = \mathbf{\Omega}^{-1}_{\boldsymbol{\varepsilon},N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}_{m,N})]\mathbf{H}_N$,

$\mathbf{F}^{**}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)]\mathbf{H}^*_N = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}_{m,N})]\mathbf{H}_N$,

where $\mathbf{H}_N$ is a $N \times P_*$ matrix whose elements are bounded uniformly in absolute value by some constant $c < \infty$. The corresponding estimates $\tilde{\mathbf{F}}_{v,N}$, $\tilde{\mathbf{F}}_{\mu,N}$, $\tilde{\mathbf{F}}^{**}_{v,N}$, and $\tilde{\mathbf{F}}^{**}_{\mu,N}$ are defined analogously, replacing σ_v^2, σ_1^2, $\mathbf{\Omega}_{\boldsymbol{\varepsilon},N}$, and $\rho_{m,N}$, $m = 1,...,S$, with $\tilde{\sigma}_v^2, \tilde{\sigma}_1^2$, $\tilde{\mathbf{\Omega}}_{\boldsymbol{\varepsilon},N}$, and $\tilde{\rho}_{m,N}$, $m = 1,...,S$, respectively.

(i) Then, $N^{-1}\tilde{\mathbf{F}}'_N\tilde{\mathbf{F}}_N - N^{-1}\mathbf{F}'_N\mathbf{F}_N = o_p(1)$ and $N^{-1}\mathbf{F}'_N\mathbf{F}_N = O(1)$.

(ii) Let $\mathbf{a}_N$ be some $NT \times 1$ or $N \times 1$ vector, whose elements are bounded uniformly in absolute value. Then $N^{-1}\tilde{\mathbf{F}}'_N\mathbf{a}_N - N^{-1}\mathbf{F}'_N\mathbf{a}_N = o_p(1)$ and $N^{-1}\mathbf{F}'_{v,N}\mathbf{a}_N = O(1)$.

Proof of part (i) of Lemma B.2

Under the maintained assumptions there exists a ρ_* with $\sup \sum_{m=1}^{S} |\rho_{m,N}| < \rho_* < 1$. It follows immediately by the properties of the matrices $\mathbf{M}_{m,N}$ that the row and column sums of $\rho_* \mathbf{M}_{m,N}$, $m = 1,...,S$ are bounded uniformly in absolute value by 1 and some finite constant respectively. For later reference, note that the elements of the vector $\rho_*^k \mathbf{M}_N^k \mathbf{h}_{.s,N}$ are also bounded uniformly in absolute value by *c* for some finite integer *k*.

Next define $\mathbf{F}_N = \mathbf{G}_N\mathbf{K}_N$ with $\mathbf{K}_N = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N} \mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$. Denote the (*r*,*s*)-th element of the difference $N^{-1}\tilde{\mathbf{F}}'_N\tilde{\mathbf{F}}_N - N^{-1}\mathbf{F}_N\mathbf{F}_N$ as v_N, which is given by

$$v_N = N^{-1}(\tilde{\mathbf{f}}'_{.r,N}\tilde{\mathbf{f}}_{.s,N} - \mathbf{f}'_{.r,N}\mathbf{f}_{.s,N}),\ r,s = 1,...,P_*, \tag{B.64a}$$

or

$$v_N = N^{-1}(\tilde{\mathbf{k}}'_{.r,N}\tilde{\mathbf{G}}'_N\tilde{\mathbf{G}}_N\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}'_{.r,N}\mathbf{G}'_N\mathbf{G}_N\mathbf{k}_{.s,N}). \tag{B.64b}$$

Define further $\mathbf{E}_N = \mathbf{G}'_N\mathbf{G}_N$, such that $\nu_N = N^{-1}(\tilde{\mathbf{k}}'_{.r,N}\tilde{\mathbf{E}}_N\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}'_{.r,N}\mathbf{E}_N\mathbf{k}_{.s,N})$.

Proof under Assumption (a)

Consider first the case $\mathbf{F}_N = \mathbf{F}_{\mu,N}$; then it holds that $\mathbf{E}_N = \tilde{\mathbf{E}}_N = T\mathbf{Q}_{1,N}$.[8] (The subsequent proof also covers the case $\mathbf{F}_N = \mathbf{F}_{v,N}$ with $\mathbf{E}_N = \tilde{\mathbf{E}}_N = \mathbf{I}_{NT}$.) Hence, $\nu_N = N^{-1}(\tilde{\mathbf{k}}'_{.r,N}\mathbf{E}_N\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}'_{.r,N}\mathbf{E}_N\mathbf{k}_{.s,N})$, which can be written as $\nu_N = \sum_{i=1}^{3}\nu_{i,N}$ with

$$\begin{aligned}
\nu_{1,N} &= N^{-1}(\tilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'\mathbf{E}_N(\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N}) \\
\nu_{2,N} &= N^{-1}(\tilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'\mathbf{E}_N\mathbf{k}_{.s,N} \\
\nu_{3,N} &= N^{-1}\mathbf{k}'_{.r,N}\mathbf{E}_N(\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N})
\end{aligned} \tag{B.65}$$

Note that

$$\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}'_{m,N})^{+}\mathbf{h}_{.s,N} - [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{h}_{.s,N} \tag{B.66}$$

We next show that $\nu_{i,N} = o_p(1)$, $i = 1,...,3$, invoking the following theorem (see, e.g., Resnik, 1999, p. 171): Let $(X, X_N, N \geq 1)$ be real valued random variables. Then, $X_N \xrightarrow{p} X$ if and only if each subsequence X_{N_a} contains a further subsequence $X_{N'_a}$ that converges almost surely to X.

As we show below we will be confronted with terms of the form:

$$\aleph_N^{(k,l)} = N^{-1}p_*^{l+k}\mathbf{h}'_{.r,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}'^{l}_N)\mathbf{E}_N(\mathbf{I}_T \otimes \overline{\mathbf{M}}^{k}_N)\mathbf{h}_{.s,N}, \tag{B.67}$$

where $\overline{\mathbf{M}}_N$ is a matrix, whose row and column sums are bounded uniformly in absolute value by some constant $c_{\overline{M}}$. By the properties of $\mathbf{E}_N = T\mathbf{Q}_{1,N}$, the row and column sums of the matrix $(\mathbf{I}_T \otimes \overline{\mathbf{M}}'^{l}_N)\mathbf{E}_N(\mathbf{I}_T \otimes \overline{\mathbf{M}}^{k}_N)$ are bounded uniformly in absolute value, and $\aleph_N^{(k,l)} = O(1)$ (compare Remark A.1 in the Appendix).

---

[8] For $\mathbf{F}_N = \mathbf{F}_{\mu,N}$, we have $\mathbf{G}_N = (\mathbf{e}'_T \otimes \mathbf{I}_N)$. Hence

$\mathbf{G}'_N\mathbf{G}_N = (\mathbf{e}_T \otimes \mathbf{I}_N)\ (\mathbf{e}'_T \otimes \mathbf{I}_N)$

$= (\mathbf{e}_T\mathbf{e}'_T \otimes \mathbf{I}_N) = (\mathbf{J}_T \otimes \mathbf{I}_N) = T\mathbf{Q}_{1,N}$.

Now, let the index N_a denote some subsequence. In light of the aforementioned equivalence, there exists a subsequence of this subsequence ($N_{a'}$) such that for events $\omega \in A$, with $P(A^C) = 0$, it holds that

$$\left|\tilde{\rho}_{m,N'_a}(\omega) - \rho_{m,N'_a}\right| \to 0\,,\ m = 1,...,S \tag{B.68}$$

and that for some $N'_a \geq \overline{N}_\omega$,

$$\left|\aleph^{(k,l)}_{N'_a}(\omega)\right| \leq K_\aleph \text{ for some } K_\aleph \text{ with } 0 < K_\aleph \leq \infty\text{, and} \tag{B.69}$$

$$\sum_{m=1}^{S}\left|\tilde{\rho}_{m,N'_a}(\omega)\right| \leq p_{**}\,,\ \text{where } p_{**} = \frac{\sup_N \sum_{m=1}^{S}\left|\rho_{m,N}\right| + p_*}{2} < 1. \tag{B.70}$$

In the following, assume that $N'_a \geq \overline{N}_\omega$. Since $\sum_{m=1}^{S}\left|\tilde{\rho}_{m,N'_a}(\omega)\right| < 1$, it follows from Horn and Johnson (1985, p. 301) that $(\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}_{m,N'_a})$ is invertible, such that

$$\begin{aligned}
\tilde{\mathbf{k}}_{.s,N'_a} - \mathbf{k}_{.s,N'_a} &= [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}'_{m,N'_a})^{-1} - \mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N'_a}\mathbf{M}'_{m,N'_a})^{-1}]\mathbf{h}_{.s,N'_a} \\
&= \{\mathbf{I}_T \otimes [\mathbf{I}_N + \sum_{l=1}^{\infty}\left(\sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}'_{m,N'_a}\right)^l] - \mathbf{I}_T \otimes [\mathbf{I}_N + \sum_{l=1}^{\infty}\left(\sum_{m=1}^{S}\rho_{m,N'_a}\mathbf{M}'_{m,N'_a}\right)^l]\}\mathbf{h}_{.s,N'_a} \\
&= \mathbf{I}_T \otimes \sum_{l=1}^{\infty}\left[\left(\sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}'_{m,N'_a}\right)^l - \left(\sum_{m=1}^{S}\rho_{m,N'_a}\mathbf{M}'_{m,N'_a}\right)^l\right]\mathbf{h}_{.s,N'_a}\,.
\end{aligned} \tag{B.71}$$

Substituting into ν_{1,N'_a}, we have

$$\begin{aligned}
\nu_{1,N'_a} &= N'^{-1}_a(\tilde{\mathbf{k}}_{.r,N'_a} - \mathbf{k}_{.r,N'_a})'\mathbf{E}_{N'_a}(\tilde{\mathbf{k}}_{.s,N'_a} - \mathbf{k}_{.s,N'_a}) \\
&= N'^{-1}_a\mathbf{h}'_{.r,N'_a}\mathbf{I}_T \otimes \sum_{l=1}^{\infty}\left[\left(\sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}_{m,N'_a}\right)^l - \left(\sum_{m=1}^{S}\rho_{m,N'_a}\mathbf{M}_{m,N'_a}\right)^l\right] \\
&\quad \times \mathbf{E}_{N'_a}\mathbf{I}_T \otimes \sum_{k=1}^{\infty}\left[\left(\sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}_{m,N'_a}\right)^k - \left(\sum_{m=1}^{S}\rho_{m,N'_a}\mathbf{M}_{m,N'_a}\right)^k\right]'\mathbf{h}_{.s,N'_a} \\
&= N'^{-1}_a\mathbf{h}'_{.r,N'_a}\mathbf{I}_T \otimes \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left[\left(\sum_{m=1}^{S}\tilde{\rho}_{m,N'_a}(\omega)\mathbf{M}_{m,N'_a}\right)^l - \left(\sum_{m=1}^{S}\rho_{m,N'_a}\mathbf{M}_{m,N'_a}\right)^l\right]
\end{aligned} \tag{B.72}$$

$$\times \mathbf{E}_{N'_a} \mathbf{I}_T \otimes \left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \mathbf{M}_{m,N'_a} \right)^k - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \mathbf{M}_{m,N'_a} \right)^k \right]' \mathbf{h}_{.s,N'_a},$$

where $\mathbf{E}_{N'_a} = T\mathbf{Q}_{1,N'_a}$ for $\mathbf{F}_N = \mathbf{F}_{\mu,N}$ (and $\mathbf{E}_N = \mathbf{I}_{NT}$ for $\mathbf{F}_N = \mathbf{F}_{v,N}$). A single element with index (*k*,*l*) of this infinite double sum over *k* and *l* is given by

$$N_a'^{-1} \mathbf{h}'_{.r,N'_a} \mathbf{I}_T \otimes \left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \mathbf{M}_{m,N'_a} \right)^l - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \mathbf{M}_{m,N'_a} \right)^l \right].$$

$$\times \mathbf{E}_{N'_a} \mathbf{I}_T \otimes \left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \mathbf{M}_{m,N'_a} \right)^k - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \mathbf{M}_{m,N'_a} \right)^k \right]' \mathbf{h}_{.s,N'_a}. \quad \text{(B.73)}$$

Next note that for any values of $\boldsymbol{\rho}_{N'_a}$ and any $\tilde{\boldsymbol{\rho}}_{N'_a}(\omega)$ there exist matrices $\hat{\mathbf{M}}_{N'_a}$ and $\breve{\mathbf{M}}_{N'_a}$, whose row and column sums are bounded uniformly in absolute value, such that:

$$\sum_{m=1}^{S} \rho_{m,N'_a} \mathbf{M}'_{m,N'_a} = \sum_{m=1}^{S} \rho_{m,N'_a} \hat{\mathbf{M}}_{N'_a} \text{ and } \sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \mathbf{M}'_{m,N'_a} = \sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \breve{\mathbf{M}}_{N'_a}. \quad \text{(B.74)}$$

$\breve{\mathbf{M}}_{N'_a}$ and $\hat{\mathbf{M}}_{N'_a}$ can thus be factored out of the sum, yielding

$$\left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \right)^l \hat{\mathbf{M}}^l_{N'_a} - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^l \breve{\mathbf{M}}^l_{N'_a} \right]. \quad \text{(B.75)}$$

By the same reasoning, for any values of $\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \right)$ and $\left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)$, there exists a matrix $\overline{\mathbf{M}}_{N'_a}$, whose row and column sums are bounded uniformly in absolute value, such that:

$$\left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \right)^l \hat{\mathbf{M}}^l_{N'_a} - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^l \breve{\mathbf{M}}^l_{N'_a} \right] = \left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \right)^l - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^l \right] \overline{\mathbf{M}}^l_{N'_a}. \quad \text{(B.76)}$$

Substituting $\overline{\mathbf{M}}_{N'_a}$ into ν_{1,N'_a}, we obtain

$$\nu_{1,N'_a} = N_a'^{-1} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \mathbf{h}'_{.r,N'_a} \mathbf{I}_T \otimes \left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \right)^l - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^l \right]' \overline{\mathbf{M}}'^l_{N'_a}$$

$$\times \mathbf{E}_{N'_a} \mathbf{I}_T \otimes \left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a}(\omega) \right)^k - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^k \right] \overline{\mathbf{M}}^k_{N'_a} \mathbf{h}_{.s,N'_a} .$$

Hence, we can write

$$\nu_{1,N'_a} = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} X^{(k,l)}_{N'_a}, \tag{B.77}$$

where $X^{(k,l)}_{N'_a} = a^{(k,l)}_{N'_a} \aleph^{(k,l)}_{N'_a}$ with

$$a^{(k,l)}_{N'_a} = \frac{\left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a} \right)^l - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^l \right]}{\rho^l_*} \frac{\left[\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a} \right)^k - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^k \right]}{\rho^k_*} \text{ and} \tag{B.78}$$

$$\aleph^{(k,l)}_{N'_a} = N'^{-1}_a p^{l+k}_* \mathbf{h}'_{.r,N'_a} (\mathbf{I}_T \otimes \overline{\mathbf{M}}'^l_{N'_a}) \mathbf{E}_{N'_a} (\mathbf{I}_T \otimes \overline{\mathbf{M}}^k_N) \mathbf{h}_{.s,N'_a} . \tag{B.79}$$

Note that $\left| a^{(k,l)}_{N'_a} \right| \to 0$ in light of the aforementioned results and since $\aleph^{(k,l)}_{N'_a} \leq K_{\aleph} < \infty$, it follows that $\left| X^{(k,l)}_{N'_a} \right| \to 0$. Moreover,

$$\left| a_{N'_a} \right| = \left| \frac{\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a} \right)^l - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^l}{\rho^l_*} \frac{\left(\sum_{m=1}^{S} \tilde{\rho}_{m,N'_a} \right)^l - \left(\sum_{m=1}^{S} \rho_{m,N'_a} \right)^k}{\rho^k_*} \right| \tag{B.80}$$

$$\leq 2 \left(\frac{\rho_{**}}{\rho_*} \right)^l 2 \left(\frac{\rho_{**}}{\rho_*} \right)^k = 4 \left(\frac{\rho_{**}}{\rho_*} \right)^{l+k} .$$

For $N_{a'} \geq \overline{N}_{\omega}$, $\left| \aleph^{(k,l)}_{N'_a}(\omega) \right| \leq K_{\aleph}$, such that we have

$$\left| X^{(k,l)}_{N'_a} \right| \leq B^{(l,k)} = 4 K_{\aleph} \left(\frac{\rho_{**}}{\rho_*} \right)^{l+k} . \tag{B.81}$$

Hence, there exists a dominating function $B^{(l,k)}$ for all values of *k,l*. Moreover, since $\left(\frac{\rho_{**}}{\rho_*} \right) < 1$ by construction, we also have that

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left|B^{(l,k)}\right| = \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}B^{(l,k)} < \infty \ , \tag{B.82}$$

i.e., the dominating function is integrable (summable). It follows from dominated convergence that

$$\lim_{N_a' \to \infty} \nu_{1,N_a'} = 0 \, . \tag{B.83}$$

The same holds for the $\nu_{i,N_a'}$, $i = 2,...,3$. It follows that $\nu_{i,N_a'} \to 0$ as $N_a' \to \infty$ and in light of Resnik, 1999, p. 171) that $\nu_N \to o_p(1)$.

Thus, $N^{-1}\widetilde{\mathbf{F}}_N'\widetilde{\mathbf{F}}_N - N^{-1}\mathbf{F}_N'\mathbf{F}_N = o_p(1)$. That $N^{-1}\mathbf{F}_N'\mathbf{E}_N\mathbf{F}_N = O(1)$ follows from the properties of $\mathbf{F}_N$ and $\mathbf{E}_N$ (compare Remark A.1). As already mentioned above, an analogous proof applies to the case $\mathbf{F}_N = \mathbf{F}_{v,N}$ with $\mathbf{E}_N = \widetilde{\mathbf{E}}_N = \mathbf{I}_{NT}$.

Proof under Assumption (b)

We first consider the case $\mathbf{F}_N = \mathbf{F}_{v,N}^{**}$. Then, $\mathbf{F}_N = \mathbf{G}_N\mathbf{K}_N$ with $\mathbf{G}_N = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1}$ and $\mathbf{K}_N = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$, such that $\mathbf{E}_N = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1}\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-1} = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}^{-2} = (\sigma_v^{-4}\mathbf{Q}_{0,N} + \sigma_1^{-4}\mathbf{Q}_{1,N})$ and $\widetilde{\mathbf{E}}_N = \widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\varepsilon},N}^{-2} = (\widetilde{\sigma}_v^{-4}\mathbf{Q}_{0,N} + \widetilde{\sigma}_1^{-4}\mathbf{Q}_{1,N})$. Notice that the subsequent proof also covers the case $\mathbf{F}_N = \mathbf{F}_{\mu,N}^{**}$.[9]

First, we rewrite $\nu_N = N^{-1}(\widetilde{\mathbf{k}}_{.r,N}'\widetilde{\mathbf{E}}_N\widetilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.r,N}'\mathbf{E}_N\mathbf{k}_{.s,N})$ as $\nu_N = \sum_{i=1}^{7}\nu_{i,N}$, with

$$\begin{aligned}
\nu_{1,N} &= N^{-1}(\widetilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'(\widetilde{\mathbf{E}}_N - \mathbf{E}_N)(\widetilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N}) \, , \\
\nu_{2,N} &= N^{-1}(\widetilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'(\widetilde{\mathbf{E}}_N - \mathbf{E}_N)\mathbf{k}_{.s,N} \, , \\
\nu_{3,N} &= N^{-1}\mathbf{k}_{.r,N}'(\widetilde{\mathbf{E}}_N - \mathbf{E}_N)(\widetilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N}) \, , \\
\nu_{4,N} &= N^{-1}\mathbf{k}_{.r,N}'(\widetilde{\mathbf{E}}_N - \mathbf{E}_N)\mathbf{k}_{.s,N} \, , \\
\nu_{5,N} &= N^{-1}(\widetilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'\mathbf{E}_N(\widetilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N}) \, ,
\end{aligned} \tag{B.84}$$

[9] In case $\mathbf{F}_N = \mathbf{F}_{\mu,N}^{**} = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$, we have $\mathbf{G}_N = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)]$, such that $\mathbf{E}_N = \sigma_1^{-2}(\mathbf{e}_T \otimes \mathbf{I}_N)\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N) = \sigma_1^{-4}(\mathbf{e}_T\mathbf{e}_T' \otimes \mathbf{I}_N) = \sigma_1^{-4}\mathbf{Q}_{1,N}$.

$$\nu_{6,N} = N^{-1}(\tilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'\mathbf{E}_N\mathbf{k}_{.s,N},$$
$$\nu_{7,N} = N^{-1}\mathbf{k}'_{.r,N}\mathbf{E}_N(\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N}).$$

For the sake of simplicity, we define $\mathbf{E}_N = \mathbf{E}_{0,N} + \mathbf{E}_{1,N}$, where $\mathbf{E}_{0,N} = \sigma_v^{-4}\mathbf{Q}_{0,N}$ and $\mathbf{E}_{1,N} = \sigma_1^{-4}\mathbf{Q}_{1,N}$ and consider only $\mathbf{E}_{1,N}$ in the following; it is obvious that an analogous proof applies to $\mathbf{E}_{0,N}$ and thus also $\mathbf{E}_N$. Next, consider

$$\tilde{\mathbf{k}}_{.s,N} - \mathbf{k}_{.s,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{h}_{.s,N} - [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{h}_{.s,N}. \qquad \text{(B.85)}$$

We next show that $\nu_{i,N} = o_p(1)$, $i = 1,...,7$. Consider

$$\mathbf{I}_T \otimes [\mathbf{I}_N - \sum_{m=1}^{S}\tilde{\rho}_{m,N}\mathbf{M}_{m,N}] - \mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}) = \mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\mathbf{M}_{m,N} \qquad \text{(B.86)}$$

and note that for any values of $\tilde{\rho}_{m,N}$ and $\rho_{m,N}$, there exists a matrix $\overline{\mathbf{M}}_N$, whose row and column sums are uniformly bounded in absolute value, such that

$$\mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\mathbf{M}_{m,N} = \mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\overline{\mathbf{M}}_N. \qquad \text{(B.87)}$$

Substituting $\overline{\mathbf{M}}_{N'_a}$ into (the first part of) the expression for $\nu_{1,N}$, we obtain

$$\begin{aligned}\nu_{1,N} = {} & N^{-1}\mathbf{h}'_{.r,N}[\mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\overline{\mathbf{M}}'_N]\tilde{\mathbf{E}}_{1,N}[\mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\overline{\mathbf{M}}_N]\mathbf{h}_{.s,N} \qquad \text{(B.88)}\\ & - N^{-1}\mathbf{h}'_{.r,N}[\mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\overline{\mathbf{M}}'_N]\mathbf{E}_{1,N}[\mathbf{I}_T \otimes \sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\overline{\mathbf{M}}_N]\mathbf{h}_{.s,N}.\end{aligned}$$

$$\begin{aligned}= {} & N^{-1}\left[\sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\right]^2 \mathbf{h}'_{.r,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}'_N)\tilde{\mathbf{E}}_{1,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)\mathbf{h}_{.s,N}\\ & - N^{-1}\left[\sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\right]^2 \mathbf{h}'_{.r,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}'_N)\mathbf{E}_{1,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)\mathbf{h}_{.s,N}.\end{aligned}$$

$$= \tilde{\sigma}_1^{-4}N^{-1}\left[\sum_{m=1}^{S}(\tilde{\rho}_{m,N} - \rho_{m,N})\right]^2 \mathbf{h}'_{.r,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}'_N)\mathbf{Q}_{1,N}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)\mathbf{h}_{.s,N}$$

$$-\sigma_1^{-4}N^{-1}\left[\sum_{m=1}^{S}(\tilde{\rho}_{m,N}-\rho_{m,N})\right]^2\mathbf{h}'_{.r,N}(\mathbf{I}_T\otimes\overline{\mathbf{M}}'_N)\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)\mathbf{h}_{.s,N}\,.$$

$$=(\tilde{\sigma}_1^{-4}-\sigma_1^{-4})N^{-1}\left[\sum_{m=1}^{S}(\tilde{\rho}_{m,N}-\rho_{m,N})\right]^2\mathbf{h}'_{.r,N}(\mathbf{I}_T\otimes\overline{\mathbf{M}}'_N)\mathbf{Q}_{1,N}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)\mathbf{h}_{.s,N}$$

$$=o_p(1)o_p(1)O(1)=o_p(1)\,.$$

The same holds for the ν_{i,N'_a}, $i=2,...,7$. Obviously, an analogous proof applies for $\mathbf{E}_{0,N}$ and thus also for $\mathbf{E}_N=\mathbf{E}_{0,N}+\mathbf{E}_{1,N}$. It follows that $\nu_N\to o_p(1)$. Thus, $N^{-1}\tilde{\mathbf{F}}'_N\tilde{\mathbf{E}}_N\tilde{\mathbf{F}}_N-N^{-1}\mathbf{F}'_N\mathbf{E}_N\mathbf{F}_N=o_p(1)$. That $N^{-1}\mathbf{F}'_N\mathbf{E}_N\mathbf{F}_N=O(1)$ follows from the properties the $\mathbf{F}_N$ elements of $\mathbf{E}_{1,N}$ ($\mathbf{E}_{0,N}$).

Proof of part (ii) of Lemma B.2

Denote the *r*-th element of the difference $N^{-1}\tilde{\mathbf{F}}'_N\mathbf{a}_N-N^{-1}\mathbf{F}'_N\mathbf{a}_N$ as w_N, which is given by

$$w_N=N^{-1}(\tilde{\mathbf{f}}'_{.r,N}\mathbf{a}_N-\mathbf{f}'_{.r,N}\mathbf{a}_N)\,,\ r=1,...,P_*\,, \tag{B.89}$$

or

$$w_N=N^{-1}(\tilde{\mathbf{k}}'_{.r,N}\tilde{\mathbf{G}}'_N\mathbf{a}_N-\mathbf{k}'_{.r,N}\mathbf{G}'_N\mathbf{a}_N)\,.$$

Proof under Assumption (a)

Consider first the case $\mathbf{F}_N=\mathbf{F}_{\mu,N}=(\mathbf{e}'_T\otimes\mathbf{I}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$. We then have $\mathbf{F}_N=\mathbf{G}_N\mathbf{K}_N$ with $\mathbf{K}_N=[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})^{-1}]\mathbf{H}_N$ and $\mathbf{G}_N=(\mathbf{e}'_T\otimes\mathbf{I}_N)$. Hence, $\mathbf{G}'_N=(\mathbf{e}_T\otimes\mathbf{I}_N)$. Obviously, the subsequent proof also covers the case $\mathbf{F}_N=\mathbf{F}_{v,N}$. It follows that

$$w_N=N^{-1}(\tilde{\mathbf{k}}'_{.r,N}-\mathbf{k}'_{.r,N})\overline{\overline{\mathbf{a}}}_N)\,, \tag{B.90}$$

where the elements of the vector $\overline{\overline{\mathbf{a}}}_N=(\mathbf{e}_T\otimes\mathbf{I}_N)\mathbf{a}_N$ are uniformly bounded in absolute value since the row and columns sums of $(\mathbf{e}_T\otimes\mathbf{I}_N)$ are uniformly bounded in absolute value and since the elements of $\mathbf{a}_N$ are uniformly bounded in absolute value. (See Remark A.1 in the Appendix.)

In light of the properties of $\mathbf{Q}_{0,N}$, $\mathbf{Q}_{1,N}$, and $\overline{\overline{\mathbf{a}}}_N$, it follows directly from the proof of part (i) of Lemma B.2, where we showed that $\nu_{2,N} = N^{-1}(\widetilde{\mathbf{k}}_{.r,N} - \mathbf{k}_{.r,N})'\mathbf{E}_N\mathbf{k}_{.s,N} = o_p(1)$, that $w_N = o_p(1)$ and thus $N^{-1}\widetilde{\mathbf{F}}_N'\mathbf{a}_N - N^{-1}\mathbf{F}_N'\mathbf{a}_N = o_p(1)$. That $N^{-1}\mathbf{F}_N'\mathbf{a}_N = O(1)$ follows immediately from the properties of $\mathbf{F}_N$ and $\mathbf{a}_N$ by Remark A.1.

Proof under Assumption (b)

Consider first the case $\mathbf{F}_N = \mathbf{F}_{v,N}^{**}$. Then, we have $\mathbf{F}_N = \mathbf{G}_N\mathbf{K}_N$ with $\mathbf{G}_N = \mathbf{\Omega}_{\varepsilon,N}^{-1}$ and $\mathbf{K}_N = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$. Notice that the subsequent proof also covers the case $\mathbf{F}_N = \mathbf{F}_{\mu,N}^{**}$. In that case,

$$
\begin{aligned}
w_N &= N^{-1}(\widetilde{\mathbf{k}}_{.r,N}'\widetilde{\mathbf{G}}_N'\mathbf{a}_N - \mathbf{k}_{.r,N}'\mathbf{G}_N'\mathbf{a}_N) \\
&= N^{-1}(\widetilde{\mathbf{k}}_{.r,N}' - \mathbf{k}_{.r,N}')\widetilde{\mathbf{\Omega}}_{\varepsilon,N}^{-1}\mathbf{a}_N + N^{-1}\mathbf{k}_{.r,N}'(\widetilde{\mathbf{\Omega}}_{\varepsilon,N}^{-1} - \mathbf{\Omega}_{\varepsilon,N}^{-1})\mathbf{a}_N .
\end{aligned}
\tag{B.91}
$$

Substituting $\widetilde{\mathbf{\Omega}}_{\varepsilon,N}^{-1} = \widetilde{\sigma}_v^{-2}\mathbf{Q}_{0,N} + \widetilde{\sigma}_\mu^{-2}\mathbf{Q}_{1,N}$, we have $w_N = w_N^v + w_N^1$, where

$$
\begin{aligned}
w_N^v &= N^{-1}\widetilde{\mathbf{k}}_{.r,N}'\widetilde{\sigma}_{v,N}^{-2}\mathbf{Q}_{0,N}\mathbf{a}_N - N^{-1}\mathbf{k}_{.r,N}'\sigma_v^{-2}\mathbf{Q}_{0,N}\mathbf{a}_N , \\
w_N^1 &= \widetilde{\sigma}_{1,N}^{-2}N^{-1}\widetilde{\mathbf{k}}_{.r,N}'\mathbf{Q}_{1,N}\mathbf{a}_N - \sigma_1^{-2}\mathbf{k}_{.r,N}'\mathbf{Q}_{1,N}\mathbf{a}_N .
\end{aligned}
$$

Considering w_N^v, we have $w_N^v = w_{1,N}^v + w_{2,N}^v$, where

$$
\begin{aligned}
w_{1,N}^v &= \widetilde{\sigma}_v^{-2}N^{-1}(\widetilde{\mathbf{k}}_{.r,N}' - \mathbf{k}_{.r,N}')\mathbf{Q}_{0,N}\mathbf{a}_N , \\
w_{2,N}^v &= (\widetilde{\sigma}_v^{-2} - \sigma_v^{-2})N^{-1}\mathbf{k}_{.r,N}'\mathbf{Q}_{0,N}\mathbf{a}_N .
\end{aligned}
\tag{B.92}
$$

The first term $w_{1,N}^v = o_p(1)$ by arguments analogous to the proof of Part (i) under Assumption (b) by the consistency of $\widetilde{\rho}_{s,N}$, $s = 1,...,S$. In light of the properties of $\mathbf{k}_{.r,N}$, $\mathbf{Q}_{0,N}$, and $\mathbf{a}_N$, $N^{-1}\mathbf{k}_{.r,N}'\mathbf{Q}_{0,N}\mathbf{a}_N = O(1)$ and since $(\widetilde{\sigma}_{v,N}^{-2} - \sigma_v^{-2}) = o_p(1)$, it follows that $w_{2,N}^v = o_p(1)$ by substituting $\widetilde{\mathbf{\Omega}}_{\varepsilon,N}^{-1} - \mathbf{\Omega}_{\varepsilon,N}^{-1} = (\widetilde{\sigma}_v^{-2} - \sigma_v^{-2})\mathbf{Q}_{0,N} + (\widetilde{\sigma}_\mu^{-2} - \sigma_\mu^{-2})\mathbf{Q}_{1,N}$. Analogous arguments apply to w_N^1 such that $w_N = o_p(1)$. Since this holds for all $r = 1,...,P_*$, we have that $N^{-1}\widetilde{\mathbf{F}}_{v,N}^{**\prime}\mathbf{a}_N - N^{-1}\mathbf{F}_{v,N}^{**\prime}\mathbf{a}_N = o_p(1)$. That $N^{-1}\mathbf{F}_{v,N}^{**\prime}\mathbf{a}_N = O(1)$ follows from immediately from the properties of $\mathbf{F}_{v,N}^{**}$ and $\mathbf{a}_N$ by Remark A.1. Analogous proof can be used to show that $N^{-1}\widetilde{\mathbf{F}}_{v,N}^{**\prime}\mathbf{a}_N - N^{-1}\mathbf{F}_{v,N}^{**\prime}\mathbf{a}_N = o_p(1)$ and that $N^{-1}\mathbf{F}_{v,N}'\mathbf{a}_N = O(1)$, which completes the proof.

Proof of Theorem 3

As part of proving Theorem 3 it has to be shown that $\tilde{\boldsymbol{\Psi}}_N - \boldsymbol{\Psi}_N = o_p(1)$. Observe that in light of (15), each element of $\tilde{\boldsymbol{\Psi}}_N$ and the corresponding element of $\boldsymbol{\Psi}_N$ can be written as

$$\begin{aligned}\tilde{\mathfrak{E}}^{p,q}_{r,s,N} &= \tilde{\mathfrak{E}}^{p,q,*}_{r,s,N} + \tilde{\mathfrak{E}}^{p,q,**}_{r,s,N} + \tilde{\mathfrak{E}}^{p,q,***}_{r,s,N} + \tilde{\mathfrak{E}}^{p,q,****}_{r,s,N}, \\ \mathfrak{E}^{p,q}_{r,s,N} &= \mathfrak{E}^{p,q,*}_{r,s,N} + \mathfrak{E}^{p,q,**}_{r,s,N} + \mathfrak{E}^{p,q,***}_{r,s,N} + \mathfrak{E}^{p,q,****}_{r,s,N},\end{aligned} \tag{B.93}$$

for $p,q = 1,...,4,(a,b)$, $r,s = 1,...,S+1$[10], where

$$\begin{aligned}\tilde{\mathfrak{E}}^{p,q,*}_{r,s,N} &= 2N^{-1}\tilde{\sigma}^4_{v,N} Tr(\mathbf{A}^v_{p,r,N}\mathbf{A}^v_{q,s,N}) + 2N^{-1}\tilde{\sigma}^4_{\mu,N} Tr(\mathbf{A}^\mu_{p,r,N}\mathbf{A}^\mu_{q,s,N}), \\ \tilde{\mathfrak{E}}^{p,q,**}_{r,s,N} &= \tilde{\sigma}^2_{v,N} N^{-1}\tilde{\mathbf{a}}^{v\,\prime}_{p,r,N}\,\tilde{\mathbf{a}}^v_{q,s,N} + \tilde{\sigma}^2_{\mu,N} N^{-1}\tilde{\mathbf{a}}^{\mu\,\prime}_{p,r,N}\,\tilde{\mathbf{a}}^\mu_{q,s,N} \\ \tilde{\mathfrak{E}}^{p,q,***}_{r,s,N} &= \tilde{\sigma}^{(3)}_{v,N} N^{-1}\sum_{i=1}^{NT}(\tilde{a}^v_{p,r,i,N}a^v_{q,s,ii,N} + a^v_{p,r,ii,N}\tilde{a}^v_{q,s,i,N}) + \tilde{\sigma}^{(3)}_{\mu,N} N^{-1}\sum_{i=1}^{N}(\tilde{a}^\mu_{p,r,i,N}a^\mu_{q,s,ii,N} + \tilde{a}^\mu_{p,r,ii,N}a^\mu_{q,s,i,N}), \\ \tilde{\mathfrak{E}}^{p,q,****}_{r,s,N} &= (\tilde{\sigma}^{(4)}_{v,N} - 3\tilde{\sigma}^4_{v,N})N^{-1}\sum_{i=1}^{NT} a^v_{p,r,ii,N}\tilde{a}^v_{q,s,ii,N} + (\tilde{\sigma}^{(4)}_{\mu,N} - 3\tilde{\sigma}^4_{\mu,N})N^{-1}\sum_{i=1}^{N}\tilde{a}^\mu_{p,r,ii,N}a^\mu_{q,s,ii,N},\end{aligned}$$

and

$$\begin{aligned}\mathfrak{E}^{p,q,*}_{r,s,N} &= 2N^{-1}\sigma^4_v Tr(\mathbf{A}^v_{p,r,N}\mathbf{A}^v_{q,s,N}) + 2N^{-1}\sigma^4_\mu Tr(\mathbf{A}^\mu_{p,r,N}\mathbf{A}^\mu_{q,s,N}), \\ \mathfrak{E}^{p,q,**}_{r,s,N} &= \sigma^2_v N^{-1}\mathbf{a}^{v\,\prime}_{p,r,N}\,\mathbf{a}^v_{q,s,N} + \sigma^2_\mu N^{-1}\mathbf{a}^{\mu\,\prime}_{p,r,N}\,\mathbf{a}^\mu_{q,s,N}, \\ \mathfrak{E}^{p,q,***}_{r,s,N} &= \sigma^{(3)}_v N^{-1}\sum_{i=1}^{NT}(a^v_{p,r,i,N}a^v_{q,s,ii,N} + a^v_{p,r,ii,N}a^v_{q,s,i,N}) + \sigma^{(3)}_\mu N^{-1}\sum_{i=1}^{N}(a^\mu_{p,r,i,N}a^\mu_{q,s,ii,N} + a^\mu_{p,r,ii,N}a^\mu_{q,s,i,N}), \\ \mathfrak{E}^{p,q,****}_{r,s,N} &= (\sigma^{(4)}_v - 3\sigma^4_v)N^{-1}\sum_{i=1}^{NT} a^v_{p,r,ii,N}a^v_{q,s,ii,N} + (\sigma^{(4)}_\mu - 3\sigma^4_\mu)N^{-1}\sum_{i=1}^{N} a^\mu_{p,r,ii,N}a^\mu_{q,s,ii,N}.\end{aligned}$$

To proof that $\tilde{\boldsymbol{\Psi}}_N - \boldsymbol{\Psi}_N = o_p(1)$, we show that $\tilde{\mathfrak{E}}^{p,q,*}_{r,s,N} - \mathfrak{E}^{p,q,*}_{r,s,N} = o_p(1)$, $\tilde{\mathfrak{E}}^{p,q,**}_{r,s,N} - \mathfrak{E}^{p,q,**}_{r,s,N} = o_p(1)$, $\tilde{\mathfrak{E}}^{p,q,***}_{r,s,N} - \mathfrak{E}^{p,q,***}_{r,s,N} = o_p(1)$, and $\tilde{\mathfrak{E}}^{p,q,****}_{r,s,N} - \mathfrak{E}^{p,q,****}_{r,s,N} = o_p(1)$.

i) Proof that $\tilde{\mathfrak{E}}^{p,q,*}_{r,s,N} - \mathfrak{E}^{p,q,*}_{r,s,N} = o_p(1)$

Consider

$$\tilde{\mathfrak{E}}^{p,q,*}_{r,s,N} - \mathfrak{E}^{p,q,*}_{r,s,N} = (\tilde{\sigma}^4_{v,N} - \sigma^4_v)2N^{-1}Tr(\mathbf{A}^v_{p,r,N}\mathbf{A}^v_{q,s,N}) + (\tilde{\sigma}^4_{\mu,N} - \sigma^4_\mu)2N^{-1}Tr(\mathbf{A}^\mu_{p,r,N}\mathbf{A}^\mu_{q,s,N}),$$

[10] See equations (34) in the main text for the structure of the matrix $\boldsymbol{\Psi}_N$ and the proper indexation of its elements.

and note that the row and column sums of the matrices $\mathbf{A}^{v}_{p,r,N}, \mathbf{A}^{v}_{q,s,N}$, $\mathbf{A}^{\mu}_{p,r,N}$, and $\mathbf{A}^{\mu}_{q,s,N}$ are uniformly bounded in absolute value by some constant, say K_A. It follows that

$$2N^{-1}Tr(\mathbf{A}^{v}_{p,r,N}\mathbf{A}^{v}_{q,s,N}) \leq 2N^{-1}NTK_A^2 = 2TK_A^2 = O(1),$$
$$2N^{-1}Tr(\mathbf{A}^{\mu}_{p,r,N}\mathbf{A}^{\mu}_{q,s,N}) \leq 2K_a^2 = 2N^{-1}NK_A^2 = O(1),$$

and thus $\widetilde{\mathfrak{E}}^{p,q,*}_{r,s,N} - \mathfrak{E}^{p,q,*}_{r,s,N} = o_p(1)$.

ii) Proof that $\widetilde{\mathfrak{E}}^{p,q,**}_{r,s,N} - \mathfrak{E}^{p,q,**}_{r,s,N} = o_p(1)$

Note that

$$\begin{aligned}\widetilde{\mathfrak{E}}^{p,q,**}_{r,s,N} - \mathfrak{E}^{p,q,**}_{r,s,N} &= \widetilde{\sigma}^2_{v,N}N^{-1}\widetilde{\boldsymbol{\alpha}}'_{p,r,N}\widetilde{\mathbf{P}}'_N\widetilde{\mathbf{F}}'_{v,N}\widetilde{\mathbf{F}}_{v,N}\widetilde{\mathbf{P}}_N\widetilde{\boldsymbol{\alpha}}_{q,s,N} + \widetilde{\sigma}^2_{\mu,N}N^{-1}\widetilde{\boldsymbol{\alpha}}'_{p,r,N}\widetilde{\mathbf{P}}'_N\widetilde{\mathbf{F}}'_{\mu,N}\widetilde{\mathbf{F}}_{\mu,N}\widetilde{\mathbf{P}}_N\widetilde{\boldsymbol{\alpha}}_{q,s,N} \\ &\quad - \sigma_v^2N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{v,N}\mathbf{F}_{v,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N} - \sigma_\mu^2N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{\mu,N}\mathbf{F}_{\mu,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N} \\ &= \Delta^{p,q,**,v}_{r,s,N} + \Delta^{p,q,**,\mu}_{r,s,N}, \text{ where}\end{aligned}$$

$$\Delta^{p,q,**,v}_{r,s,N} = \widetilde{\sigma}^2_{v,N}N^{-1}\widetilde{\boldsymbol{\alpha}}'_{p,r,N}\widetilde{\mathbf{P}}'_N\widetilde{\mathbf{F}}'_{v,N}\widetilde{\mathbf{F}}_{v,N}\widetilde{\mathbf{P}}_N\widetilde{\boldsymbol{\alpha}}_{q,s,N} - \sigma_v^2N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{v,N}\mathbf{F}_{v,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N}, \text{ and}$$
$$\Delta^{p,q,**,\mu}_{r,s,N} = \widetilde{\sigma}^2_{\mu,N}N^{-1}\widetilde{\boldsymbol{\alpha}}'_{p,r,N}\widetilde{\mathbf{P}}'_N\widetilde{\mathbf{F}}'_{\mu,N}\widetilde{\mathbf{F}}_{\mu,N}\widetilde{\mathbf{P}}_N\widetilde{\boldsymbol{\alpha}}_{q,s,N} - \sigma_\mu^2N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{\mu,N}\mathbf{F}_{\mu,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N}.$$

Consider $\Delta^{p,q,**,v}_{r,s,N}$, which can be written as

$$\begin{aligned}\Delta^{p,q,**,v}_{r,s,N} &= \widetilde{\sigma}^2_{v,N}N^{-1}\widetilde{\boldsymbol{\alpha}}'_{p,r,N}\widetilde{\mathbf{P}}'_N\widetilde{\mathbf{F}}'_{v,N}\widetilde{\mathbf{F}}_{v,N}\widetilde{\mathbf{P}}_N\widetilde{\boldsymbol{\alpha}}_{q,s,N} - \sigma_v^2N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{v,N}\mathbf{F}_{v,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N} \\ \Delta^{p,q,**,v}_{r,s,N} &= (\widetilde{\sigma}^2_{v,N} - \sigma_v^2)N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{v,N}\mathbf{F}_{v,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N} \\ &\quad + \widetilde{\sigma}^2_{v,N}(N^{-1}\widetilde{\boldsymbol{\alpha}}'_{p,r,N}\widetilde{\mathbf{P}}'_N\widetilde{\mathbf{F}}'_{v,N}\widetilde{\mathbf{F}}_{v,N}\widetilde{\mathbf{P}}_N\widetilde{\boldsymbol{\alpha}}_{q,s,N} - N^{-1}\boldsymbol{\alpha}'_{p,r,N}\mathbf{P}'_N\mathbf{F}'_{v,N}\mathbf{F}_{v,N}\mathbf{P}_N\boldsymbol{\alpha}_{q,s,N}).\end{aligned}$$

Note that $(\widetilde{\sigma}^2_{v,N} - \sigma_v^2) = o_p(1)$ and that $\widetilde{\sigma}^2_{v,N} = O_p(1)$. By assumption $\widetilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$, $\mathbf{P}_N = O(1)$ and thus $\widetilde{\mathbf{P}}_N = O_p(1)$, where the dimension of $\mathbf{P}_N$ is $P_* \times P$; by Lemma B.1 $\widetilde{\boldsymbol{\alpha}}_N - \boldsymbol{\alpha}_N = o_p(1)$, $\boldsymbol{\alpha}_N = O(1)$ and thus $\widetilde{\boldsymbol{\alpha}}_N = O_p(1)$, where the dimension of $\boldsymbol{\alpha}_N$ is $P \times 1$. Moreover, $N^{-1}\widetilde{\mathbf{F}}'_{v,N}\widetilde{\mathbf{F}}_{v,N} - N^{-1}\mathbf{F}'_{v,N}\mathbf{F}_{v,N}$ is $o_p(1)$ and $N^{-1}\mathbf{F}'_{v,N}\mathbf{F}_{v,N} = O(1)$ by Lemma B.2. It follows that $\Delta^{p,q,**,v}_{r,s,N} = o_p(1)$. By Lemma B.2 it also holds that $N^{-1}\widetilde{\mathbf{F}}'_{\mu,N}\widetilde{\mathbf{F}}_{\mu,N} - N^{-1}\mathbf{F}'_{\mu,N}\mathbf{F}_{\mu,N}$ and accordingly, $\Delta^{p,q,**,\mu}_{r,s,N} = o_p(1)$. It follows that $\mathfrak{E}^{p,q,**}_{r,s,N} - \widetilde{\mathfrak{E}}^{p,q,**}_{r,s,N} = o_p(1)$, and in light of Lemma B.2

it is readily observed that this also holds when $\mathbf{F}^{**}_{\mu,N}$ and $\mathbf{F}^{**}_{v,N}$ are used instead of $\mathbf{F}_{\mu,N}$ and $\mathbf{F}_{v,N}$.

iii) Proof that $\widetilde{\mathfrak{E}}^{p,q,***}_{r,s,N} - \mathfrak{E}^{p,q,***}_{r,s,N} = o_p(1)$

Observe that

$$\widetilde{\mathfrak{E}}^{p,q,***}_{r,s,N} - \mathfrak{E}^{p,q,***}_{r,s,N} = \Delta^{p,q,***,v}_{r,s,N} + \Delta^{p,q,***,\mu}_{r,s,N},$$

where

$$\Delta^{p,q,***,v}_{r,s,N} = \widetilde{\sigma}^{(3)}_{v,N} N^{-1} \sum_{i=1}^{NT} (\widetilde{a}^{v}_{p,r,i,N} a^{v}_{q,s,ii,N} + a^{v}_{p,r,ii,N} \widetilde{a}^{v}_{q,s,i,N}) - \sigma^{(3)}_{v,N} N^{-1} \sum_{i=1}^{NT} (a^{v}_{p,r,i,N} a^{v}_{q,s,ii,N} + a^{v}_{p,r,ii,N} a^{v}_{q,s,i,N}),$$

$$\Delta^{p,q,***,\mu}_{r,s,N} = \widetilde{\sigma}^{(3)}_{\mu,N} N^{-1} \sum_{i=1}^{N} (\widetilde{a}^{\mu}_{p,r,i,N} a^{\mu}_{q,s,ii,N} + a^{\mu}_{p,r,ii,N} \widetilde{a}^{\mu}_{q,s,i,N}) - \sigma^{(3)}_{\mu,N} N^{-1} \sum_{i=1}^{N} (a^{\mu}_{p,r,i,N} a^{\mu}_{q,s,ii,N} + a^{\mu}_{p,r,ii,N} a^{\mu}_{q,s,i,N}).$$

Next consider $\Delta^{p,q,***,v}_{r,s,N}$ and note that $\widetilde{\mathbf{a}}^{v}_{p,r,N} = \widetilde{\mathbf{F}}_{v,N} \widetilde{\mathbf{P}}_N \widetilde{\boldsymbol{\alpha}}_{p,r,N}$; moreover, define $\bar{\mathbf{a}}^{v}_{p,r,N}$ as $NT \times 1$ vector made up of the main diagonal elements $a^{v}_{p,r,ii,N}$ of matrix $\mathbf{A}^{v}_{p,r,N}$. It follows that

$$\begin{aligned}\Delta^{p,q,***,v}_{r,s,N} &= \widetilde{\sigma}^{(3)}_{v,N} N^{-1} (\widetilde{\mathbf{a}}^{v\,\prime}_{p,r,i,N} \bar{\mathbf{a}}^{v}_{q,s,N} - \mathbf{a}^{v\,\prime}_{p,r,N} \bar{\mathbf{a}}^{v}_{q,s,N}) + \widetilde{\sigma}^{(3)}_{v,N} N^{-1} (\bar{\mathbf{a}}^{v\,\prime}_{p,r,N} \widetilde{\mathbf{a}}^{v}_{q,s,N} - \bar{\mathbf{a}}^{v\,\prime}_{p,r,N} \mathbf{a}^{v}_{q,s,N}) \\ &\quad + (\widetilde{\sigma}^{(3)}_{v,N} - \sigma^{(3)}_{v,N}) N^{-1} (\mathbf{a}^{v\,\prime}_{p,r,N} \bar{\mathbf{a}}^{v}_{q,s,N} + \bar{\mathbf{a}}^{v\,\prime}_{p,r,N} \mathbf{a}^{v}_{q,s,N}).\end{aligned}$$

Note that $\sigma^{(3)}_{v} = O(1)$, $(\widetilde{\sigma}^{(3)}_{v,N} - \sigma^{(3)}_{v}) = o_p(1)$, $\widetilde{\sigma}^{(3)}_{v,N} = O_p(1)$. Moreover, by the properties of $\mathbf{a}^{v\,\prime}_{p,r,N}$, $\bar{\mathbf{a}}^{v}_{p,r,N}$, $N^{-1}(\mathbf{a}^{v\,\prime}_{p,r,N} \bar{\mathbf{a}}^{v}_{q,s,N} + \bar{\mathbf{a}}^{v\,\prime}_{p,r,N} \mathbf{a}^{v}_{q,s,N}) = O(1)$ such that the term in the second line is $o_p(1)$. Next note that $\widetilde{\mathbf{a}}^{v\,\prime}_{p,r,N} \bar{\mathbf{a}}^{v}_{p,r,N} = \widetilde{\boldsymbol{\alpha}}'_{p,r,N} \widetilde{\mathbf{P}}'_N \widetilde{\mathbf{F}}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N}$, and $\mathbf{a}^{v\,\prime}_{p,r,N} \bar{\mathbf{a}}^{v}_{p,r,N} = \boldsymbol{\alpha}'_{p,r,N} \mathbf{P}'_N \mathbf{F}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N}$, such that the first term in the first line is given by $\widetilde{\boldsymbol{\alpha}}'_{p,r,N} \widetilde{\mathbf{P}}'_N N^{-1} \widetilde{\mathbf{F}}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N} - \boldsymbol{\alpha}'_{p,r,N} \mathbf{P}'_N N^{-1} \mathbf{F}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N}$. By assumption $\widetilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$, $\mathbf{P}_N = O(1)$ and thus $\widetilde{\mathbf{P}}_N = O_p(1)$, where the dimension of $\mathbf{P}_N$ is $P_* \times P$. Moreover, $\widetilde{\boldsymbol{\alpha}}_N - \boldsymbol{\alpha}_N = o_p(1)$, $\boldsymbol{\alpha}_N = O(1)$ and thus $\widetilde{\boldsymbol{\alpha}}_N = O_p(1)$, where the dimension of $\boldsymbol{\alpha}_N$ is $P \times 1$. By Lemma B.2, part (ii), we have that $N^{-1} \widetilde{\mathbf{F}}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N} - N^{-1} \mathbf{F}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N} = o_p(1)$ and $N^{-1} \mathbf{F}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N} = O(1)$. It follows that $\widetilde{\boldsymbol{\alpha}}'_{p,r,N} \widetilde{\mathbf{P}}'_N N^{-1} \widetilde{\mathbf{F}}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N} - \boldsymbol{\alpha}'_{p,r,N} \mathbf{P}'_N N^{-1} \mathbf{F}'_{v,N} \bar{\mathbf{a}}^{v}_{p,r,N} = o_p(1)$. By analogous arguments the second term in the first line is $o_p(1)$, from which it follows that $\Delta^{p,q,***,v}_{r,s,N} = o_p(1)$. An analogous proof can be used to show that $\Delta^{p,q,***,\mu}_{r,s,N} = o_p(1)$. Hence, $\widetilde{\mathfrak{E}}^{p,q,***}_{r,s,N} - \mathfrak{E}^{p,q,***}_{r,s,N} = o_p(1)$.

iv) Proof that $\widetilde{\mathfrak{E}}^{p,q,****}_{r,s,N} - \mathfrak{E}^{p,q,****}_{r,s,N} = o_p(1)$

Consider

$$\widetilde{\mathfrak{E}}^{p,q,****}_{r,s,N} - \mathfrak{E}^{p,q,****}_{r,s,N} = [(\widetilde{\sigma}^{(4)}_v - 3\widetilde{\sigma}^4_v) - (\sigma^{(4)}_v - 3\sigma^4_v)]N^{-1}\overline{\mathbf{a}}^{v\,\prime}_{p,r,N}\,\overline{\mathbf{a}}^{v}_{q,s,N} + [(\widetilde{\sigma}^{(4)}_\mu - 3\widetilde{\sigma}^4_\mu) - (\sigma^{(4)}_\mu - 3\sigma^4_\mu)]N^{-1}\overline{\mathbf{a}}^{\mu\,\prime}_{p,r,N}\,\overline{\mathbf{a}}^{\mu}_{q,s,N}.$$

By the properties of the matrices $\mathbf{A}^{v}_{p,r,N}, \mathbf{A}^{v}_{q,s,N}$, $\mathbf{A}^{\mu}_{p,r,N}$, and $\mathbf{A}^{\mu}_{q,s,N}$, it follows by arguments analogous to above that $N^{-1}\overline{\mathbf{a}}^{v\,\prime}_{p,r,N}\,\overline{\mathbf{a}}^{v}_{q,s,N} = O(1)$ and that $N^{-1}\overline{\mathbf{a}}^{\mu\,\prime}_{p,r,N}\,\overline{\mathbf{a}}^{\mu}_{q,s,N} = O(1)$. Since $(\widetilde{\sigma}^{(4)}_{v,N} - \sigma^{(4)}_v) = o_p(1)$ and $(\widetilde{\sigma}^{(4)}_{\mu,N} - \sigma^{(4)}_\mu) = o_p(1)$, and since $(\widetilde{\sigma}^2_{v,N} - \sigma^2_v) = o_p(1)$, $(\widetilde{\sigma}^2_{\mu,N} - \sigma^2_\mu) = o_p(1)$, and thus also $(\widetilde{\sigma}^4_{v,N} - \sigma^4_v) = o_p(1)$, $(\widetilde{\sigma}^4_{\mu,N} - \sigma^4_\mu) = o_p(1)$, it follows that $\widetilde{\mathfrak{E}}^{p,q,****}_{r,s,N} - \mathfrak{E}^{p,q,****}_{r,s,N} = o_p(1)$.

We have shown that $\widetilde{\mathfrak{E}}^{p,q,*}_{r,s,N} - \mathfrak{E}^{p,q,*}_{r,s,N} = o_p(1)$, $\widetilde{\mathfrak{E}}^{p,q,**}_{r,s,N} - \mathfrak{E}^{p,q,**}_{r,s,N} = o_p(1)$, $\widetilde{\mathfrak{E}}^{p,q,***}_{r,s,N} - \mathfrak{E}^{p,q,***}_{r,s,N} = o_p(1)$, and $\widetilde{\mathfrak{E}}^{p,q,****}_{r,s,N} - \mathfrak{E}^{p,q,****}_{r,s,N} = o_p(1)$ for all elements of $\mathbf{\Psi}_N$, $p,q = 1,...,4,(a,b)$, $r,s = 1,...,S+1$. It follows that $\widetilde{\mathbf{\Psi}}_N - \mathbf{\Psi}_N = o_p(1)$.

We are now ready to prove consistency of the estimate of the variance-covariance matrix, i.e., that $\widetilde{\mathbf{\Omega}}_{\tilde{\boldsymbol{\theta}}_N} - \mathbf{\Omega}_{\tilde{\boldsymbol{\theta}}_N} = o_p(1)$, where $\mathbf{\Omega}_{\tilde{\boldsymbol{\theta}}_N}(\mathbf{\Theta}_N) = (\mathbf{J}'_N\mathbf{\Theta}_N\mathbf{J}_N)^{-1}\mathbf{J}'_N\mathbf{\Theta}_N\mathbf{\Psi}_N\mathbf{\Theta}_N\mathbf{J}_N(\mathbf{J}'_N\mathbf{\Theta}_N\mathbf{J}_N)^{-1}$. As proved above, we have $\widetilde{\mathbf{\Psi}}_N - \mathbf{\Psi}_N = o_p(1)$. By Assumption 5, we have $\widetilde{\mathbf{\Theta}}_N - \mathbf{\Theta}_N = o_p(1)$, $\mathbf{\Theta}_N = O(1)$ and $\widetilde{\mathbf{\Theta}}_N = O_p(1)$. Let $\mathbf{\Xi}_N = \mathbf{J}'_N\mathbf{\Theta}\mathbf{J}_N = \mathfrak{B}'_N\,\mathbf{\Gamma}'_N\,\mathbf{\Theta}_N\mathbf{\Gamma}_N\mathfrak{B}_N$ as defined in Theorem 2 and $\widetilde{\mathbf{\Xi}}_N = \widetilde{\mathbf{J}}'_N\widetilde{\mathbf{\Theta}}\widetilde{\mathbf{J}}_N = \widetilde{\mathfrak{B}}'_N\,\widetilde{\mathbf{\Gamma}}'_N\widetilde{\mathbf{\Theta}}_N\widetilde{\mathbf{\Gamma}}_N\widetilde{\mathfrak{B}}_N$.[11] In Theorem 2, we showed that $\widetilde{\mathbf{J}}_N = O_p(1)$, $\mathbf{J}_N = O(1)$, and $\widetilde{\mathbf{J}}_N - \mathbf{J}_N = o_p(1)$ and that $\widetilde{\mathbf{\Xi}}^{+}_N = O_p(1)$, $\mathbf{\Xi}^{-1}_N = O_p(1)$ and that $\widetilde{\mathbf{\Xi}}^{+}_N - \mathbf{\Xi}^{-1}_N = o_p(1)$. It now follows that $\widetilde{\mathbf{\Omega}}_{\tilde{\boldsymbol{\theta}}_N} - \mathbf{\Omega}_{\tilde{\boldsymbol{\theta}}_N} = o_p(1)$.

III. Proof of Theorem 4 (Joint Distribution of $\tilde{\boldsymbol{\rho}}_N$ and Other Model Parameters)

[11] There is a slight discrepancy to the definition of $\widetilde{\mathbf{\Xi}}_N$ in Theorem 2: Here $\widetilde{\mathfrak{B}}_N$ is used rather than $\overline{\mathfrak{B}}_N$, which does not affect the proof, however, noting that both $\tilde{\boldsymbol{\rho}}_N$ and $\bar{\boldsymbol{\rho}}_N$ are consistent.

The subsequent proof will focus on the case $\mathbf{F}_{v,N}$ and $\mathbf{F}_{v,N}$; this also convers the case for $\mathbf{F}_{v,N}^{**}$ and $\mathbf{F}_{\mu,N}^{**}$. The first line in Theorem 4 holds in light of Assumption 7 (for $N^{1/2}\boldsymbol{\Delta}_N$), bearing in mind that $\mathbf{T}_{v,N} = \mathbf{F}_{v,N}\mathbf{P}_N$, and Theorem 2 (for $N^{1/2}(\tilde{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_N)$).

We next prove that $\boldsymbol{\xi}_{o,N} = \boldsymbol{\Psi}_{o,N}^{-1/2}(N^{-1/2}\boldsymbol{\xi}_N'\mathbf{F}_N, \mathbf{q}_N')' \xrightarrow{d} N(\mathbf{0}, \mathbf{I}_{P^*+4S+2})$ by verifying that the assumptions of the central limit theorem A.1 by Kelejian and Prucha (2008) are fulfilled. Note that $\lambda_{\min}(\boldsymbol{\Psi}_{o,N}) \geq c_{\boldsymbol{\Psi}_o}^* > 0$ by assumption. In Theorem 2, we verified that the stacked innovations $\boldsymbol{\xi}_N$, the matrices $\mathbf{A}_{1,s,N},...,\mathbf{A}_{4,s,N}$, $s = 1,...,S$, $\mathbf{A}_{a,N}$, and $\mathbf{A}_{b,N}$, and the vectors $\mathbf{a}_{1,s,N}$, …, $\mathbf{a}_{4,s,N}$, $s = 1,...,S$, $\mathbf{a}_{a,N}$, and $\mathbf{a}_{b,N}$ satisfy the assumptions of central limit theorem by Kelejian and Prucha (2008, Theorem A.1).

Next, consider the two blocks of $\mathbf{F}_N = (\mathbf{F}_{v,N}', \mathbf{F}_{\mu,N}')'$, which are given by

$$\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N}')^{-1}]\mathbf{H}_N, \text{ and}$$

$$\mathbf{F}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)]\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N}')^{-1}]\mathbf{H}_N.$$

Since the row and columns sums of $[\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)]$, $\boldsymbol{\Omega}_{\boldsymbol{\varepsilon},N}$, and $[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S} \rho_{m,N}\mathbf{M}_{m,N}')^{-1}]$ are uniformly bounded in absolute value and since the elements of the matrix $\mathbf{H}_N$ are uniformly bounded in absolute value, it follows that the elements of $\mathbf{F}_N$ are also uniformly bounded in absolute value. Hence, the linear form $\mathbf{F}_N'\boldsymbol{\xi}_N = \mathbf{F}_{v,N}'\mathbf{v}_N + \mathbf{F}_{\mu,N}'\boldsymbol{\mu}_N$ also fulfils the assumptions of Theorem A.1. As a consequence, $\boldsymbol{\xi}_{o,N} \xrightarrow{d} N(\mathbf{0}, \mathbf{I}_{P^*+4S+2})$.

In the proofs of Theorems 2 and 3, we showed that $\tilde{\boldsymbol{\Psi}}_N - \boldsymbol{\Psi}_N = o_p(1)$, $\boldsymbol{\Psi}_N = O(1)$, and $\tilde{\boldsymbol{\Psi}}_N = O_p(1)$. By analogous arguments, this also holds for the submatrices $\boldsymbol{\Psi}_{\Delta\Delta,N}$ and $\boldsymbol{\Psi}_{\Delta\theta,N}$. Hence, $\tilde{\boldsymbol{\Psi}}_{o,N} - \boldsymbol{\Psi}_{o,N} = o_p(1)$, $\boldsymbol{\Psi}_{o,N} = O(1)$ and $\tilde{\boldsymbol{\Psi}}_{o,N} - \boldsymbol{\Psi}_{o,N} = o_p(1)$, and thus $\tilde{\boldsymbol{\Psi}}_{o,N} = O_p(1)$.

By assumption $\tilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$, $\mathbf{P}_N = O(1)$, and $\tilde{\mathbf{P}}_N = O_p(1)$ as well as $\tilde{\boldsymbol{\Theta}}_N - \boldsymbol{\Theta}_N = o_p(1)$, $\boldsymbol{\Theta}_N = O(1)$ and $\tilde{\boldsymbol{\Theta}}_N = O_p(1)$. In the proof of Theorem 2 we showed that $\tilde{\mathbf{J}}_N - \mathbf{J}_N = o_p(1)$, $\mathbf{J}_N = O(1)$, and $\tilde{\mathbf{J}}_N = O_p(1)$, and that $(\tilde{\mathbf{J}}_N'\tilde{\boldsymbol{\Theta}}_N\tilde{\mathbf{J}}_N)^+ - (\mathbf{J}_N'\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1} = o_p(1)$, $(\mathbf{J}_N'\boldsymbol{\Theta}_N\mathbf{J}_N)^{-1} = O(1)$,

and $(\widetilde{\mathbf{J}}_N'\widetilde{\mathbf{\Theta}}_N\widetilde{\mathbf{J}}_N)^+ = O_p(1)$. It now follows that $\widetilde{\mathbf{\Omega}}_{o,N} - \mathbf{\Omega}_{o,N} = o_p(1)$ and $\mathbf{\Omega}_{o,N} = O(1)$ and thus $\widetilde{\mathbf{\Omega}}_{o,N} = O_p(1)$.

APPENDIX C.

Proof of Lemma 1.

In light of equations (4a) and (4b), Assumptions 3 and 8, as well as $\sup_N \|\boldsymbol{\beta}_N\| \leq b < \infty$, it follows that all columns of $\mathbf{Z}_N = (\mathbf{X}_N, \overline{\mathbf{Y}}_N)$ are of the form $\vartheta_N = \boldsymbol{\pi}_N + \mathbf{\Pi}_N\boldsymbol{\varepsilon}_N$, where the elements of the vector $\boldsymbol{\pi}_N$ and the row and column sums of the matrix $\mathbf{\Pi}_N$ are bounded uniformly in absolute value (see Remark A.1 in Appendix A). It follows from Lemma C.2 in Kelejian and Prucha (2008) that the fourth moments of the elements of the matrix $\mathbf{D}_N = -\mathbf{Z}_N$ are bounded uniformly by some finite constant and that Assumption 6 holds.

Next, note that

$$(NT)^{1/2}(\widetilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N) = \widetilde{\mathbf{P}}_N'(NT)^{-1/2}\mathbf{F}_{v,N}'\mathbf{v}_N + \widetilde{\mathbf{P}}_N'(NT)^{-1/2}\mathbf{F}_{\mu,N}'\boldsymbol{\mu}_N,$$

where $\widetilde{\mathbf{P}}_N$ is defined in the Lemma, and

$$\mathbf{F}_{v,N} = [\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')^{-1}]\mathbf{H}_N \text{ and}$$

$$\mathbf{F}_{\mu,N} = [\sigma_1^{-2}(\mathbf{e}_T' \otimes \mathbf{I}_N)]\mathbf{\Omega}_{\varepsilon,N}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')^{-1}]\mathbf{H}_N.$$

In light of Assumption 10, $\widetilde{\mathbf{P}}_N - \mathbf{P}_N = o_p(1)$ and $\mathbf{P}_N = O(1)$, with $\mathbf{P}_N$ as defined in the Lemma. By Assumptions 2, 3 and 9, the elements of $\mathbf{F}_{v,N}$ and $\mathbf{F}_{\mu,N}$ are bounded uniformly in absolute value. By Assumption 2, $E(\mathbf{v}_N) = \mathbf{0}$, $E(\boldsymbol{\mu}_N) = \mathbf{0}$, and the diagonal variance-covariance matrices of $\mathbf{v}_N$ and $\boldsymbol{\mu}_N$ have uniformly bounded elements. Thus, $E[(NT)^{-1/2}\mathbf{F}_{v,N}'\mathbf{v}_N] = \mathbf{0}$ and the elements of the variance-covariance matrix of $N^{-1/2}\mathbf{F}_{v,N}'\mathbf{v}_N$, i.e., $(NT)^{-1}\sigma_v^2\mathbf{F}_{v,N}'\mathbf{F}_{v,N}$, are bounded uniformly in absolute value (see Remark A.1 in Appendix A). Moreover, $E[(NT)^{-1/2}\mathbf{F}_{2,N}'\boldsymbol{\mu}_N] = \mathbf{0}$, and the elements of the variance-covariance matrix of $N^{-1/2}\mathbf{F}_{\mu,N}'\boldsymbol{\mu}_N$, i.e., $(NT)^{-1}\sigma_\mu^2\mathbf{F}_{\mu,N}'\mathbf{F}_{\mu,N}$, are bounded uniformly in absolute value. It follows from Chebychev's inequality that $(NT)^{-1/2}\mathbf{F}_{v,N}'\mathbf{v}_N = O_p(1)$, $(NT)^{-1/2}\mathbf{F}_{\mu,N}'\boldsymbol{\mu}_N = O_p(1)$, and consequently $(NT)^{1/2}(\widetilde{\boldsymbol{\delta}}_N - \boldsymbol{\delta}_N) = \mathbf{P}_N(NT)^{-1/2}\mathbf{F}_{v,N}'\mathbf{v}_N + \mathbf{P}_N(NT)^{-1/2}\mathbf{F}_{\mu,N}'\boldsymbol{\mu}_N + o_p(1)$ and

$\mathbf{P}_N(NT)^{-1/2}\mathbf{F}'_{v,N}\mathbf{v}_N + \mathbf{P}_N(NT)^{-1/2}\mathbf{F}'_{\mu,N}\boldsymbol{\mu}_N = O_p(1)$. This completes the proof, recalling that $\mathbf{T}_N = (\mathbf{T}'_{v,N}, \mathbf{T}'_{\mu,N})' = (\mathbf{P}'_N\mathbf{F}'_{v,N}, \mathbf{P}'_N\mathbf{F}'_{\mu,N})'$.

Proof of Lemma 2.
The FGTSLS estimator is given by

$$\hat{\breve{\boldsymbol{\delta}}}_N = (\hat{\breve{\mathbf{Z}}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**})^{-1}\hat{\breve{\mathbf{Z}}}_N^{**\prime}\breve{\mathbf{y}}_N^{**}, \text{ where } \breve{\mathbf{y}}_N^{**} = \breve{\mathbf{Z}}_N^{**}\boldsymbol{\delta}_N + \breve{\mathbf{u}}_N^{**} \text{ with}$$

$$\breve{\mathbf{u}}_N^{**} = \breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\breve{\rho}_{m,N}\mathbf{M}_{m,N})]\mathbf{u}_N .$$

Substituting $\hat{\breve{\mathbf{Z}}}_N^{**} = \breve{\mathbf{P}}_{\mathbf{H}_N^{**}}\breve{\mathbf{Z}}_N^{**} = \breve{\mathbf{H}}_N^{**}(\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{H}}_N^{**})^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**}$, we obtain

$$(NT)^{1/2}[\hat{\breve{\boldsymbol{\delta}}}_N - \boldsymbol{\delta}_N] = (NT)^{1/2}\boldsymbol{\Delta}_N^{**} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{u}}_N^{**}, \text{ with}$$

$$\breve{\mathbf{P}}_N^{**\prime} = \{[(NT)^{-1}\breve{\mathbf{Z}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}][(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}]^{-1}[(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**}]\}^{-1}[(NT)^{-1}\breve{\mathbf{Z}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}][(NT)^{-1}\breve{\mathbf{H}}_N^{**\prime}\breve{\mathbf{H}}_N^{**}]^{-1}.$$

Next note that

$$\begin{aligned}\mathbf{u}_N^{**}(\breve{\boldsymbol{\theta}}_N) = \mathbf{u}_N^{**}(\boldsymbol{\theta}_N) &- \sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)\mathbf{u}_N ,\\ &+ (\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2} - \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{u}_N\\ &+ (\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2} - \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)\mathbf{u}_N\\ &- (\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2} - \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)\mathbf{u}_N ,\end{aligned}$$

where $\overline{\mathbf{M}}_N$ is a matrix, whose row and columns sums are bounded uniformly in absolute value, satisfying

$$\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\mathbf{M}_{m,N} = \sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\overline{\mathbf{M}}_N .$$

Substituting for $\breve{\mathbf{u}}_N^{**}$, we obtain

$$(NT)^{1/2}(\hat{\breve{\boldsymbol{\delta}}}_N - \boldsymbol{\delta}_N) = (NT)^{1/2}\boldsymbol{\Delta}_N^{**} = \boldsymbol{\vartheta}_{1,N} + \boldsymbol{\vartheta}_{2,N} + \boldsymbol{\vartheta}_{3,N} + \boldsymbol{\vartheta}_{4,N} + \boldsymbol{\vartheta}_{5,N} = \sum_{i=1}^{5}\boldsymbol{\vartheta}_{i,N}, \text{ where}$$

$$\mathfrak{d}_{1,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\breve{\mathbf{H}}_N^{**\prime}\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{2,N} = -(NT)^{-1/2}\mathbf{P}_N^{**}(\breve{\boldsymbol{\theta}}_N)'\breve{\mathbf{H}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{3,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\breve{\mathbf{H}}_N^{**\prime}(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\,\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{4,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\breve{\mathbf{H}}_N^{**\prime}(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{5,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\breve{\mathbf{H}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N.$$

Note that the FGTSLS estimator uses (generated) transformed instruments, which are based on the estimate $\hat{\boldsymbol{\theta}}_N$. Observe that

$$\begin{aligned}\breve{\mathbf{H}}_N^{**} = \mathbf{H}_N^{**} &- \sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)\mathbf{H}_N\\ &+(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N\\ &+(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)\mathbf{H}_N\\ &-(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)\mathbf{H}_N.\end{aligned}$$

Substituting for $\breve{\mathbf{H}}_N^{**\prime}$, we obtain $\mathfrak{d}_{i,N}=\sum_{j=1}^{5}\mathfrak{d}_{ij,N}$, $i=1,...,5$. Considering $\mathfrak{d}_{1,N}$, we have

$$\mathfrak{d}_{11,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{F}_{1,N}^{**\prime}\mathbf{v}_N + (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{F}_{2,N}^{**\prime}\boldsymbol{\mu}_N, \text{ with}$$

$$\mathbf{F}_{1,N}^{**\prime} = \mathbf{H}_N^{*\prime}[(\sigma_v^{-2}\mathbf{Q}_{0,N}+\sigma_1^{-2}\mathbf{Q}_{1,N}) = \mathbf{H}_N^{*\prime}\boldsymbol{\Omega}_{\varepsilon,N}^{-1}, \text{ and}$$

$$\mathbf{F}_{2,N}^{**\prime} = \mathbf{H}_N^{*\prime}\sigma_1^{-2}(\mathbf{e}_T\otimes\mathbf{I}_N).$$

$$\mathfrak{d}_{12,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}'_N(\mathbf{I}_T\otimes\overline{\mathbf{M}}'_N)\boldsymbol{\Omega}_{\varepsilon,N}^{-1}\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{13,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}'_N[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}'_{m,N})](\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{14,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}'_N(\mathbf{I}_T\otimes\overline{\mathbf{M}}'_N)(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{15,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}'_N(\mathbf{I}_T\otimes\overline{\mathbf{M}}'_N)(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}\boldsymbol{\varepsilon}_N$$

Regarding $\mathfrak{d}_{2,N}$ we have

$$\mathfrak{d}_{21,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N^{**\prime}\sum_{m=1}^{S}(\hat{\rho}_{m,N}-\rho_{m,N})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{22,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}[\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})]^2\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')\boldsymbol{\Omega}_{\varepsilon,N}^{-1}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{23,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}_N'[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')]$$
$$\times(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{24,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')$$
$$\times(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{25,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}[\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})]^2\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')$$
$$\times(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N)[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N.$$

Regarding $\mathfrak{d}_{3,N}$ we have

$$\mathfrak{d}_{31,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N^{**\prime}(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{32,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{33,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N'[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')](\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\varepsilon}_N$$
$$= (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N'[\mathbf{I}_T\otimes(\mathbf{I}_N-\sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}')](\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})^2\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{34,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\varepsilon}_N$$
$$= (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})^2\boldsymbol{\varepsilon}_N,$$

$$\mathfrak{d}_{35,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})\boldsymbol{\varepsilon}_N$$
$$= -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N}-\rho_{m,N})\mathbf{H}_N'(\mathbf{I}_T\otimes\overline{\mathbf{M}}_N')(\breve{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2}-\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2})^2\boldsymbol{\varepsilon}_N.$$

Regarding $\mathfrak{d}_{4,N}$ we have

$$\mathfrak{d}_{41,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N^{**\prime}(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N ,$$

$$\mathfrak{d}_{42,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\mathbf{H}_N^{\prime}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N^{\prime})\mathbf{\Omega}_{\varepsilon,N}^{-1/2}$$
$$\times(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N ,$$

$$\mathfrak{d}_{43,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N^{\prime}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}^{\prime})]$$
$$\times(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})^2(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N$$

$$\mathfrak{d}_{44,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime} + \mathbf{H}_N^{\prime}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N^{\prime})(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})^2(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N$$

$$\mathfrak{d}_{45,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\mathbf{H}_N^{\prime}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N^{\prime})$$
$$\times(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})^2(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N$$

Regarding $\mathfrak{d}_{5,N}$ we have

$$\mathfrak{d}_{51,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{H}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N ,$$

$$\mathfrak{d}_{52,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}[\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})]^2\mathbf{H}_N^{\prime}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N^{\prime})\mathbf{\Omega}_{\varepsilon,N}^{-1/2}$$
$$\times(\hat{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N ,$$

$$\mathfrak{d}_{53,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\mathbf{H}_N^{\prime}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N}^{\prime})]$$
$$\times(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})^2(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N ,$$

$$\mathfrak{d}_{54,N} = -(NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})\mathbf{H}_N^{\prime}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N^{\prime})$$
$$\times(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})^2(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N ,$$

$$\mathfrak{d}_{55,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}[\sum_{m=1}^{S}(\breve{\rho}_{m,N} - \rho_{m,N})]^2\mathbf{H}_N^{\prime}(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N^{\prime})$$
$$\times(\breve{\mathbf{\Omega}}_{\varepsilon,N}^{-1/2} - \mathbf{\Omega}_{\varepsilon,N}^{-1/2})^2(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]\boldsymbol{\varepsilon}_N .$$

Summing up we have

$$(NT)^{1/2}\boldsymbol{\Delta}_N^{**} = \sum_{i=1}^{5}\sum_{j=1}^{5}\boldsymbol{\mathfrak{d}}_{ij,N}\ .$$

Next note that, in light of Assumption 10 and since $\breve{\boldsymbol{\theta}}_N$ is $N^{1/2}$-consistent, it follows that

$$(NT)^{-1}\hat{\tilde{\mathbf{Z}}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**} - \mathbf{Q}'_{\mathbf{H}^{**}\mathbf{Z}^{**}}\mathbf{Q}^{-1}_{\mathbf{H}^{**}\mathbf{H}^{**}}\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}} = o_p(1)\,. \tag{D.4}$$

By Assumption 10b we also have $\mathbf{Q}'_{\mathbf{H}^{**}\mathbf{Z}^{**}}\mathbf{Q}^{-1}_{\mathbf{H}^{**}\mathbf{H}^{**}}\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}} = O(1)$ and thus $(\mathbf{Q}'_{\mathbf{H}^{**}\mathbf{Z}^{**}}\mathbf{Q}^{-1}_{\mathbf{H}^{**}\mathbf{H}^{**}}\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}})^{-1} = O(1)$. It follows as a special case of Pötscher and Prucha (1997, Lemma F1) that

$$[\,(NT)^{-1}\hat{\tilde{\mathbf{Z}}}_N^{**\prime}\breve{\mathbf{Z}}_N^{**}\,]^{-1} - (\mathbf{Q}'_{\mathbf{H}^{**}\mathbf{Z}^{**}}\mathbf{Q}^{-1}_{\mathbf{H}^{**}\mathbf{H}^{**}}\mathbf{Q}_{\mathbf{H}^{**}\mathbf{Z}^{**}})^{-1} = o_p(1)\,. \tag{D.5}$$

It follows further that $\breve{\mathbf{P}}_N^{**} - \mathbf{P}_N^{**} = o_p(1)$ and $\mathbf{P}_N^{**} = O(1)$ with $\mathbf{P}_N^{**}$ defined in the Lemma.

Next observe that $(\breve{\boldsymbol{\rho}}_N - \boldsymbol{\rho}_N) = o_p(1)$ and that $(\breve{\boldsymbol{\Omega}}_{\varepsilon,N} - \boldsymbol{\Omega}_{\varepsilon,N}) = o_p(1)\mathbf{I}_{NT}$. Note further that all terms $\boldsymbol{\mathfrak{d}}_{ij,N}$ except for $\boldsymbol{\mathfrak{d}}_{11,N}$ are of the form $o_p(1)\breve{\mathbf{P}}_N^{**\prime}(NT)^{-1/2}\boldsymbol{\mathfrak{D}}'_N\boldsymbol{\varepsilon}_N$, where $\boldsymbol{\mathfrak{D}}_N$ are $NT \times P_*$ matrices involving products of $(\mathbf{I}_T \otimes \overline{\mathbf{M}}_N)$, $[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})^{-1}]$, $\mathbf{H}_N$, $\boldsymbol{\Omega}_{\varepsilon,N}$ ($\boldsymbol{\Omega}_{\varepsilon,N}^{-1}$), and $\boldsymbol{\Omega}_{\varepsilon,N}^{1/2}$ ($\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}$). By the maintained assumptions regarding these matrices it follows that the elements of $\boldsymbol{\mathfrak{D}}_N$ are bounded uniformly in absolute value. As a consequence, $E[(NT)^{-1/2}\boldsymbol{\mathfrak{D}}'_N\boldsymbol{\varepsilon}_N] = \mathbf{0}$ and the elements of the variance-covariance matrix of $(NT)^{-1/2}\boldsymbol{\mathfrak{D}}'_N\boldsymbol{\varepsilon}_N$, i.e., $(NT)^{-1}\boldsymbol{\mathfrak{D}}'_N\boldsymbol{\Omega}_{\varepsilon,N}\boldsymbol{\mathfrak{D}}_N$, are bounded uniformly in absolute value (see Remark A.1 in Appendix A). It follows from Chebychev's inequality that $(NT)^{-1/2}\boldsymbol{\mathfrak{D}}'_N\boldsymbol{\varepsilon}_N = O_p(1)$. As a consequence, all terms $\boldsymbol{\mathfrak{d}}_{ij,N}$ except for $\boldsymbol{\mathfrak{d}}_{11,N}$ are $o_p(1)$, and $\boldsymbol{\mathfrak{d}}_{11,N} = O_p(1)$. Finally, observe that $\boldsymbol{\mathfrak{d}}_{11,N} = (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{F}_{v,N}^{**\prime}\mathbf{v}_N + (NT)^{-1/2}\breve{\mathbf{P}}_N^{**\prime}\mathbf{F}_{\mu,N}^{**\prime}\boldsymbol{\mu}_N$, with

$\mathbf{F}_{v,N}^{**} = (\sigma_v^{-2}\mathbf{Q}_{0,N} + \sigma_1^{-2}\mathbf{Q}_{1,N})\mathbf{H}_N^* = \boldsymbol{\Omega}_{\varepsilon,N}^{-1}[\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$, and

$\mathbf{F}_{\mu,N}^{**} = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)]\mathbf{H}_N^* = [\sigma_1^{-2}(\mathbf{e}'_T \otimes \mathbf{I}_N)][\mathbf{I}_T \otimes (\mathbf{I}_N - \sum_{m=1}^{S}\rho_{m,N}\mathbf{M}_{m,N})]\mathbf{H}_N$,

which completes the proof, recalling that $\mathbf{T}_N^{**} = \mathbf{F}_N^{**}\mathbf{P}_N^{**}$.

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2493 Christian Bruns and Oliver Himmler, It's the Media, Stupid – How Media Activity Shapes Public Spending, December 2008

2494 Andreas Knabe and Ronnie Schöb, Minimum Wages and their Alternatives: A Critical Assessment, December 2008

2495 Sascha O. Becker, Peter H. Egger, Maximilian von Ehrlich and Robert Fenge, Going NUTS: The Effect of EU Structural Funds on Regional Performance, December 2008

2496 Robert Dur, Gift Exchange in the Workplace: Money or Attention?, December 2008

2497 Scott Alan Carson, Nineteenth Century Black and White US Statures: The Primary Sources of Vitamin D and their Relationship with Height, December 2008

2498 Thomas Crossley and Mario Jametti, Pension Benefit Insurance and Pension Plan Portfolio Choice, December 2008

2499 Sebastian Hauptmeier, Ferdinand Mittermaier and Johannes Rincke, Fiscal Competition over Taxes and Public Inputs: Theory and Evidence, December 2008

2500 Dirk Niepelt, Debt Maturity without Commitment, December 2008

2501 Andrew Clark, Andreas Knabe and Steffen Rätzel, Boon or Bane? Others' Unemployment, Well-being and Job Insecurity, December 2008

2502 Lukas Menkhoff, Rafael R. Rebitzky and Michael Schröder, Heterogeneity in Exchange Rate Expectations: Evidence on the Chartist-Fundamentalist Approach, December 2008

2503 Salvador Barrios, Harry Huizinga, Luc Laeven and Gaëtan Nicodème, International Taxation and Multinational Firm Location Decisions, December 2008

2504 Andreas Irmen, Cross-Country Income Differences and Technology Diffusion in a Competitive World, December 2008

2505 Wenan Fei, Claude Fluet and Harris Schlesinger, Uncertain Bequest Needs and Long-Term Insurance Contracts, December 2008

2506 Wido Geis, Silke Uebelmesser and Martin Werding, How do Migrants Choose their Destination Country? An Analysis of Institutional Determinants, December 2008

2507 Hiroyuki Kasahara and Katsumi Shimotsu, Sequential Estimation of Structural Models with a Fixed Point Constraint, December 2008

2508 Barbara Hofmann, Work Incentives? Ex Post Effects of Unemployment Insurance Sanctions – Evidence from West Germany, December 2008

2509 Louis Hotte and Stanley L. Winer, The Demands for Environmental Regulation and for Trade in the Presence of Private Mitigation, December 2008

2510 Konstantinos Angelopoulos, Jim Malley and Apostolis Philippopoulos, Welfare Implications of Public Education Spending Rules, December 2008

2511 Robert Orlowski and Regina T. Riphahn, The East German Wage Structure after Transition, December 2008

2512 Michel Beine, Frédéric Docquier and Maurice Schiff, International Migration, Transfers of Norms and Home Country Fertility, December 2008

2513 Dirk Schindler and Benjamin Weigert, Educational and Wage Risk: Social Insurance vs. Quality of Education, December 2008

2514 Bernd Hayo and Stefan Voigt, The Relevance of Judicial Procedure for Economic Growth, December 2008

2515 Bruno S. Frey and Susanne Neckermann, Awards in Economics – Towards a New Field of Inquiry, January 2009

2516 Gregory Gilpin and Michael Kaganovich, The Quantity and Quality of Teachers: A Dynamic Trade-off, January 2009

2517 Sascha O. Becker, Peter H. Egger and Valeria Merlo, How Low Business Tax Rates Attract Multinational Headquarters: Municipality-Level Evidence from Germany, January 2009

2518 Geir H. Bjønnes, Steinar Holden, Dagfinn Rime and Haakon O.Aa. Solheim, ‚Large' vs. ‚Small' Players: A Closer Look at the Dynamics of Speculative Attacks, January 2009

2519 Jesus Crespo Cuaresma, Gernot Doppelhofer and Martin Feldkircher, The Determinants of Economic Growth in European Regions, January 2009

2520 Salvador Valdés-Prieto, The 2008 Chilean Reform to First-Pillar Pensions, January 2009

2521 Geir B. Asheim and Tapan Mitra, Sustainability and Discounted Utilitarianism in Models of Economic Growth, January 2009

2522 Etienne Farvaque and Gaël Lagadec, Electoral Control when Policies are for Sale, January 2009

2523 Nicholas Barr and Peter Diamond, Reforming Pensions, January 2009

2524 Eric A. Hanushek and Ludger Woessmann, Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation, January 2009

2525 Richard Arnott and Eren Inci, The Stability of Downtown Parking and Traffic Congestion, January 2009

2526 John Whalley, Jun Yu and Shunming Zhang, Trade Retaliation in a Monetary-Trade Model, January 2009

2527 Mathias Hoffmann and Thomas Nitschka, Securitization of Mortgage Debt, Asset Prices and International Risk Sharing, January 2009

2528 Steven Brakman and Harry Garretsen, Trade and Geography: Paul Krugman and the 2008 Nobel Prize in Economics, January 2009

2529 Bas Jacobs, Dirk Schindler and Hongyan Yang, Optimal Taxation of Risky Human Capital, January 2009

2530 Annette Alstadsæter and Erik Fjærli, Neutral Taxation of Shareholder Income? Corporate Responses to an Announced Dividend Tax, January 2009

2531 Bruno S. Frey and Susanne Neckermann, Academics Appreciate Awards – A New Aspect of Incentives in Research, January 2009

2532 Nannette Lindenberg and Frank Westermann, Common Trends and Common Cycles among Interest Rates of the G7-Countries, January 2009

2533 Erkki Koskela and Jan König, The Role of Profit Sharing in a Dual Labour Market with Flexible Outsourcing, January 2009

2534 Tomasz Michalak, Jacob Engwerda and Joseph Plasmans, Strategic Interactions between Fiscal and Monetary Authorities in a Multi-Country New-Keynesian Model of a Monetary Union, January 2009

2535 Michael Overesch and Johannes Rincke, What Drives Corporate Tax Rates Down? A Reassessment of Globalization, Tax Competition, and Dynamic Adjustment to Shocks, February 2009

2536 Xenia Matschke and Anja Schöttner, Antidumping as Strategic Trade Policy Under Asymmetric Information, February 2009

2537 John Whalley, Weimin Zhou and Xiaopeng An, Chinese Experience with Global 3G Standard-Setting, February 2009

2538 Claus Thustrup Kreiner and Nicolaj Verdelin, Optimal Provision of Public Goods: A Synthesis, February 2009

2539 Jerome L. Stein, Application of Stochastic Optimal Control to Financial Market Debt Crises, February 2009

2540 Lars P. Feld and Jost H. Heckemeyer, FDI and Taxation: A Meta-Study, February 2009

2541 Philipp C. Bauer and Regina T. Riphahn, Age at School Entry and Intergenerational Educational Mobility, February 2009

2542 Thomas Eichner and Rüdiger Pethig, Carbon Leakage, the Green Paradox and Perfect Future Markets, February 2009

2543 M. Hashem Pesaran, Andreas Pick and Allan Timmermann, Variable Selection and Inference for Multi-period Forecasting Problems, February 2009

2544 Mathias Hoffmann and Iryna Shcherbakova, Consumption Risk Sharing over the Business Cycle: the Role of Small Firms' Access to Credit Markets, February 2009

2545 John Beirne, Guglielmo Maria Caporale, Marianne Schulze-Ghattas and Nicola Spagnolo, Volatility Spillovers and Contagion from Mature to Emerging Stock Markets, February 2009

2546 Ali Bayar and Bram Smeets, Economic and Political Determinants of Budget Deficits in the European Union: A Dynamic Random Coefficient Approach, February 2009

2547 Jan K. Brueckner and Anming Zhang, Airline Emission Charges: Effects on Airfares, Service Quality, and Aircraft Design, February 2009

2548 Dolores Messer and Stefan C. Wolter, Money Matters – Evidence from a Large-Scale Randomized Field Experiment with Vouchers for Adult Training, February 2009

2549 Johannes Rincke and Christian Traxler, Deterrence through Word of Mouth, February 2009

2550 Gabriella Legrenzi, Asymmetric and Non-Linear Adjustments in Local Fiscal Policy, February 2009

2551 Bruno S. Frey, David A. Savage and Benno Torgler, Surviving the Titanic Disaster: Economic, Natural and Social Determinants, February 2009

2552 Per Engström, Patrik Hesselius and Bertil Holmlund, Vacancy Referrals, Job Search, and the Duration of Unemployment: A Randomized Experiment, February 2009

2553 Giorgio Bellettini, Carlotta Berti Ceroni and Giovanni Prarolo, Political Persistence, Connections and Economic Growth, February 2009

2554 Steinar Holden and Fredrik Wulfsberg, Wage Rigidity, Institutions, and Inflation, February 2009

2555 Alexander Haupt and Tim Krieger, The Role of Mobility in Tax and Subsidy Competition, February 2009

2556 Harald Badinger and Peter Egger, Estimation of Higher-Order Spatial Autoregressive Panel Data Error Component Models, February 2009