Genuinely Unbalanced Spatial Panel Data Models with Fixed Effects: M-Estimation and Inference with an Application to FDI Inflows

Xiaoyu Meng[1] and Zhenlin Yang[2]

[1]*Nankai University and* [2]*Singapore Management University*

Supplementary Material

This `Supplementary Material` contains additional appendices: Appendix B presents some basic lemmas, Appendix C proofs of the main results in Section 3, Appendix D proofs of the main results in Section 4, and Appendix E a complete set of Monte Carlo results.

Appendix B: Some Basic Lemmas

The following lemmas are essential to the proofs of the main results in this paper. Lemmas B.1, B.2 and B.5 are taken from the literature and thus their proofs are omitted.

Lemma B.1. *(Kelejian and Prucha, 1999; Lee, 2002): Let $\{A_N\}$ and $\{B_N\}$ be two sequences of $N \times N$ matrices that are bounded in both row and column sum norms. Let C_N be a sequence of conformable matrices whose elements are uniformly $O(h_n^{-1})$. Then,*

(i) the sequence $\{A_N B_N\}$ are uniformly bounded in both row and column sums,

(ii) the elements of A_N are uniformly bounded and $tr(A_N) = O(N)$, and

(iii) the elements of $A_N C_N$ and $C_N A_N$ are uniformly $O(h_n^{-1})$.

Lemma B.2. *(Lemma A.3, Lee, 2004): For $\mathbf{W}$ and $\mathbf{A}_N(\lambda)$ defined in Model (3.1), if $\|\mathbf{W}\|$*

and $\|\mathbf{A}_N^{-1}\|$ *are uniformly bounded, where* $\|\cdot\|$ *is a matrix norm, then* $\|\mathbf{A}_N^{-1}(\lambda)\|$ *is uniformly bounded in a neighborhood of* λ_0.

Lemma B.3. *Under Assumptions C-E and H, we have*

(*i*) $\mathbb{Q}_{\mathbb{D}}(\rho)$ *is bounded in both row and column sum norms, uniformly in* $\rho \in \Delta_\rho$*;*

(*ii*) $\mathbb{Q}_{\mathbb{X}}(\rho)$ *is bounded in both row and column sum norms, uniformly in* $\rho \in \Delta_\rho$*.*

Proof of Lemma B.3: Proof is simpler using a $\mathbf{D}_\alpha^\star$ under the constraint $\alpha_1 = 0$.

Proof of (*i*)**.** Recall $\mathbb{D}(\rho) = [\mathbb{D}_\mu(\rho), \mathbb{D}_\alpha(\rho)]$ with $\mathbb{D}_\mu(\rho) = \mathbf{B}_N(\rho)\mathbf{D}_\mu$ and $\mathbb{D}_\alpha(\rho) = \mathbf{B}_N(\rho)\mathbf{D}_\alpha^\star$. Denote $\mathcal{D}_{11}(\rho) = \mathbb{D}_\mu'(\rho)\mathbb{D}_\mu(\rho)$, $\mathcal{D}_{12}(\rho) = \mathbb{D}_\mu'(\rho)\mathbb{D}_\alpha(\rho)$, $\mathcal{D}_{22}(\rho) = \mathbb{D}_\alpha'(\rho)\mathbb{D}_\alpha(\rho)$. Using the inverse formula of a partitioned matrix, we have

$$[\mathbb{D}'(\rho)\mathbb{D}(\rho)]^{-1} = \begin{bmatrix} \mathcal{F}^{-1}(\rho) & -\mathcal{F}^{-1}(\rho)\mathcal{D}_{12}(\rho)\mathcal{D}_{22}^{-1}(\rho) \\ -\mathcal{D}_{22}^{-1}(\rho)\mathcal{D}_{12}'(\rho)\mathcal{F}^{-1}(\rho) & \mathcal{D}_{22}^{-1}(\rho) + \mathcal{D}_{22}^{-1}(\rho)\mathcal{D}_{12}'(\rho)\mathcal{F}^{-1}(\rho)\mathcal{D}_{12}(\rho)\mathcal{D}_{22}^{-1}(\rho) \end{bmatrix},$$

where $\mathcal{F}(\rho) = \mathcal{D}_{11}(\rho) - \mathcal{D}_{12}(\rho)\mathcal{D}_{22}^{-1}(\rho)\mathcal{D}_{12}'(\rho)$. Plugging this into $\mathbb{Q}_{\mathbb{D}}(\rho)$, we obtain,

$$\mathbb{Q}_{\mathbb{D}}(\rho) = \mathbb{Q}_{\mathbb{D}_\alpha}(\rho) - \mathbb{Q}_{\mathbb{D}_\alpha}(\rho)\mathbb{D}_\mu(\rho)[\mathbb{D}_\mu'(\rho)\mathbb{Q}_{\mathbb{D}_\alpha}(\rho)\mathbb{D}_\mu(\rho)]^{-1}\mathbb{D}_\mu'(\rho)\mathbb{Q}_{\mathbb{D}_\alpha}(\rho). \tag{B.1}$$

Given the special structure of $\mathbb{D}_\alpha(\rho)$, one has $\mathbb{Q}_{\mathbb{D}_\alpha}(\rho) = \texttt{blkdiag}(J_1(\rho), \ldots, J_T(\rho))$, where $J_1(\rho) = I_{n_1}$ and $J_t(\rho) = I_{n_t} - \frac{1}{n_t}B_t(\rho)l_{n_t}[\frac{1}{n_t}l_{n_t}'B_t'(\rho)B_t(\rho)l_{n_t}]^{-1}l_{n_t}'B_t'(\rho)$ for $t = 2, \cdots, T$. By Assumption D, the limit of $\frac{1}{n_t}l_{n_t}'B_t'(\rho)B_t(\rho)l_{n_t}$ is bounded away from zero and the elements of $B_t(\rho)l_{n_t}l_{n_t}'B_t'(\rho)$ are uniformly bounded, uniformly in $\rho \in \Delta_\rho$ for each t. Therefore, $J_t(\rho)$ must be uniformly bounded in both row and column sums, uniformly in $\rho \in \Delta_\rho$, $\forall t$. Hence, $\mathbb{Q}_{\mathbb{D}_\alpha}(\rho)$ is also uniformly bounded in both row and column sums, uniformly in $\rho \in \Delta_\rho$.

We next consider the second term on the RHS of equation (B.1). We denote it as $\bar{\mathcal{Q}}(\rho)$, which can be partitioned into $T \times T$ blocks with (s,t)th block being

$$\bar{\mathcal{Q}}_{s,t}(\rho) = -\tfrac{1}{T}J_s(\rho)B_s(\rho)D_s[\tfrac{1}{T}\textstyle\sum_{t=1}^T D_t'B_t'(\rho)J_t(\rho)B_t(\rho)D_t]^{-1}D_t'B_t'(\rho)J_t(\rho).$$

By assuming $B_s(\rho)D_s[\frac{1}{T}\sum_{t=1}^{T} D_t' B_t'(\rho) J_t(\rho) B_t(\rho) D_t]^{-1} D_t' B_t'(\rho)$ is uniformly bounded in both row and column sum norms, uniformly in $\rho \in \Delta_\rho$, for all s and t, we have that the row and column sums of each $\bar{\mathcal{Q}}_{s,t}(\rho)$ must have uniform order $O(1/T)$, uniformly in $\rho \in \Delta_\rho$. As there are T blocks in each row or in each column of $\bar{\mathcal{Q}}(\rho)$, we must have $\bar{\mathcal{Q}}(\rho)$ is bounded in both row and column sum norms, uniformly in $\rho \in \Delta_\rho$. Consequently, $\mathbb{Q}_{\mathbb{D}}(\rho)$ is bounded in both row and column sum norms, uniformly in $\rho \in \Delta_\rho$.

Proof of (ii)**.** Let $Z_N(\rho) = [\frac{1}{N}\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}$ with its (j,k)th element being denoted by $z_{jk}(\rho)$. From Assumption C(ii), $Z_N(\rho)$ converges to a finite limit uniformly in $\rho \in \Delta_\rho$. Therefore, there exists a constant c_z such that $|z_{jk}(\rho)| \leq c_z$ uniformly in $\rho \in \Delta_\rho$ for large enough N. Note that $\mathbb{X}(\rho) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)\mathbf{X}$. As the elements of $\mathbf{X}$ are uniformly bounded (Assumption C(i)), and $\mathbf{B}_N(\rho)$ and $\mathbb{Q}_{\mathbb{D}}(\rho)$ are bounded in both row and column sum norms, uniformly in $\rho \in \Delta_\rho$, the elements of $\mathbb{X}(\rho)$ are also uniformly bounded, uniformly in $\rho \in \Delta_\rho$. Hence, there exists a constant c_x such that $|x_{jk}(\rho)| \leq c_x$ uniformly in $\rho \in \Delta_\rho$, where $x_{jk}(\rho)$ is the (j,k)th element of $\mathbb{X}(\rho)$. Let $p_{jl}(\rho)$ be the (j,l)th element of $\mathbb{P}_{\mathbb{X}}(\rho) = \frac{1}{N}\mathbb{X}(\rho)[\frac{1}{N}\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}\mathbb{X}'(\rho)$. It follows that uniformly in $\rho \in \Delta_\rho$, $\sum_{j=1}^{N}|p_{jl}(\rho)| \leq \frac{1}{N}\sum_{j=1}^{N}\sum_{r=1}^{k}\sum_{s=1}^{k}|z_{rs}(\rho)x_{jr}(\rho)x_{ls}(\rho)| \leq k^2 c_z c_x^2$ for all $l = 1, 2, \ldots, N$. Similarly, uniformly in $\rho \in \Delta_\rho$, we have $\sum_{l=1}^{N}|p_{jl}(\rho)| \leq \frac{1}{N}\sum_{l=1}^{N}\sum_{r=1}^{k}\sum_{s=1}^{k}|z_{rs}(\rho)x_{jr}(\rho)x_{ls}(\rho)| \leq k^2 c_z c_x^2$ for all $j = 1, 2, \ldots, N$. That is, $\|\mathbb{P}_{\mathbb{X}}(\rho)\|_1$ and $\|\mathbb{P}_{\mathbb{X}}(\rho)\|_\infty$ are bounded, uniformly in $\rho \in \Delta_\rho$. Consequently, $\|\mathbb{Q}_{\mathbb{X}}(\rho)\|_1$ and $\|\mathbb{Q}_{\mathbb{X}}(\rho)\|_\infty$ are bounded, uniformly in $\rho \in \Delta_\rho$. □

Lemma B.4. *Suppose that* $\{A_N\}$ *and* $\{B_N\}$ *are two sequences of* $N \times N$ *matrices that are uniformly bounded in either row or column sums. Under Assumptions C-E and H,* $\mathtt{tr}[A_N \mathbb{P}_{\mathbb{X}}(\rho) B_N] = O(1)$*, uniformly in* $\rho \in \Delta_\rho$*.*

Proof of Lemma B.4: From the proof of Lemma B.3, the elements of $\mathbb{X}(\rho)$ and the elements of $[\frac{1}{N}\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}$ are uniformly bounded, uniformly in $\rho \in \Delta_\rho$. If A_N and B_N are bounded in row (column) sum norm, then $A_N B_N$ is also bounded in row (column) sum norm. Thus, Lemma A.6 of Lee (2004) implies that the elements of $\frac{1}{N}\mathbb{X}'(\rho)A_N B_N \mathbb{X}(\rho)$ are uniformly bounded. It follows $\mathtt{tr}[A_N \mathbb{P}_{\mathbb{X}}(\rho) B_N] = \mathtt{tr}[(\frac{1}{N}\mathbb{X}'(\rho)\mathbb{X}(\rho))^{-1}\frac{1}{N}\mathbb{X}'(\rho)A_N B_N \mathbb{X}(\rho)] = O(1)$, uniformly in $\rho \in \Delta_\rho$ because the number of regressors k is fixed. □

Lemma B.5. *(Lemma A.2, Lin and Lee, 2010; Lemma A.3, Liu and Yang, 2015): Let $A_N = [a_{ij}]$ and $B_N = [b_{ij}]$ be two square matrices of dimension N and c_N be an $N \times 1$ vector of elements c_i. Assume that innovations $\{v_j\}$ are independent with zero mean, i.e. $v_j \sim inid(0, \sigma_j^2)$. Letting $\mathbf{H} = \mathtt{diag}\{\sigma_1^2, \cdots, \sigma_N^2\}$ and $\mathbf{V} = (v_1, \cdots, v_N)'$, we have,*

(*i*) $\mathrm{E}(\mathbf{V}'A_N\mathbf{V}) = \mathtt{tr}(\mathbf{H}A_N) = \sum_{i=1}^N a_{ii}\sigma_i^2$,

(*ii*) $\mathrm{E}(\mathbf{V}'A_N\mathbf{V} \cdot c_N'\mathbf{V}) = \sum_{i=1}^N a_{ii}c_i \mathrm{E}(v_i^3)$,

(*iii*) $\mathrm{E}(\mathbf{V}'A_N\mathbf{V} \cdot \mathbf{V}'B_N\mathbf{V}) = \sum_{i=1}^N a_{ii}b_{ii}[\mathrm{E}(v_i^4) - 3\sigma_i^4] + \mathtt{tr}(\mathbf{H}A_N)\mathtt{tr}(\mathbf{H}B_N) + \mathtt{tr}(\mathbf{H}A_N\mathbf{H}B_N^\circ)$,

(*iv*) $\mathrm{Var}(\mathbf{V}'A_N\mathbf{V}) = \sum_{i=1}^N a_{ii}^2[\mathrm{E}(v_i^4) - 3\sigma_i^4] + \mathtt{tr}(\mathbf{H}A_N\mathbf{H}A_N^\circ)$.

Lemma B.6. *(Lemma A.3, Lin and Lee, 2010, extended): Let $\{A_N\}$ be a sequence of $N \times N$ matrices such that either $\|A_N\|_\infty$ or $\|A_N\|_1$ is bounded. Suppose that the elements of A_N are $O(h_n^{-1})$ uniformly in all i and j. Let innovation vector $\mathbf{V}$ be defined as in Lemma B.5. Let c_N be an $N \times 1$ vector with elements of uniform order $O(h_n^{-1/2})$. Then*

(*i*) $\mathrm{E}(\mathbf{V}'A_N\mathbf{V}) = O(\frac{N}{h_n})$, (*ii*) $\mathrm{Var}(\mathbf{V}'A_N\mathbf{V}) = O(\frac{N}{h_n})$,

(*iii*) $\mathbf{V}'A_N\mathbf{V} = O_p(\frac{N}{h_n})$, (*iv*) $\mathbf{V}'A_N\mathbf{V} - \mathrm{E}(\mathbf{V}'A_N\mathbf{V}) = O_p((\frac{N}{h_n})^{\frac{1}{2}})$,

(*v*) $c_N'A_N\mathbf{V} = O_p((\frac{N}{h_n})^{\frac{1}{2}})$, **if** $\|A_N\|_1$ *is bounded.*

Proof of Lemma B.6: Firstly, Lemma A.8 of Lee (2004) implies that $\mathtt{tr}(\mathbf{H}A_N)$, $\mathtt{tr}(A_N A_N')$, $\mathtt{tr}(\mathbf{H}A_N\mathbf{H}A_N)$ and $\mathtt{tr}(\mathbf{H}A_N\mathbf{H}A_N')$ are all $O(\frac{N}{h_n})$. As $\sum_{i=1}^N a_{ii}^2 \leq \mathtt{tr}(A_N A_N')$, we also have

$\sum_{i=1}^{N} a_{ii}^2 = O(\frac{N}{h_n})$. These and Lemma B.5 show that $\mathrm{E}(\mathbf{V}'A_N\mathbf{V}) = \mathtt{tr}(\mathbf{H}A_N) = O(\frac{N}{h_n})$ and $\mathrm{Var}(\mathbf{V}'A_N\mathbf{V}) = \sum_{i=1}^{N} a_{ii}^2[\mathrm{E}(v_i^4) - 3\sigma_i^4] + \mathtt{tr}[\mathbf{H}A_N(\mathbf{H}A_N' + \mathbf{H}A_N)] = O(\frac{N}{h_n})$. As $\mathrm{E}[(\mathbf{V}'A_N\mathbf{V})^2] = \mathrm{Var}(\mathbf{V}'A_N\mathbf{V}) + \mathrm{E}^2(\mathbf{V}'A_N\mathbf{V}) = O((\frac{N}{h_n})^2)$, we have $P(\frac{h_n}{N}|\mathbf{V}'A_N\mathbf{V}| \geq M) \leq \frac{1}{M^2}(\frac{h_n}{N})^2\mathrm{E}[(\mathbf{V}'A_N\mathbf{V})^2] = O(1)$, by the generalized Chebyshev's inequality. It follows that $\mathbf{V}'A_N\mathbf{V} = O_p(\frac{N}{h_n})$. Moreover, by Chebyshev's inequality, $P((\frac{h_n}{N})^{\frac{1}{2}}|\mathbf{V}'A_N\mathbf{V} - \mathrm{E}(\mathbf{V}'A_N\mathbf{V})| \geq M) \leq \frac{1}{M^2}\frac{h_n}{N}\mathrm{Var}(\mathbf{V}'A_N\mathbf{V}) = O(1)$. This implies that $\mathbf{V}'A_N\mathbf{V} - \mathrm{E}(\mathbf{V}'A_N\mathbf{V}) = O_p((\frac{N}{h_n})^{\frac{1}{2}})$. Finally, as the elements of c_N have uniform order $O(h_n^{-1/2})$, there exists a constant $\bar{c}$ such that $|c_j| \leq \frac{\bar{c}}{h_n^{1/2}}$ for all j. Hence, by the boundedness of $\|A_N\|_1$,

$$\begin{aligned}\mathrm{Var}[(\tfrac{h_n}{N})^{\frac{1}{2}}c_N'A_N\mathbf{V}] &= \tfrac{h_n}{N}\textstyle\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} c_jc_ka_{ji}a_{ki}\sigma_i^2 \\ &\leq \bar{c}^2(\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\sigma_i^2)(\sum_{j=1}^{N}|a_{ji}|)(\sum_{k=1}^{N}|a_{ki}|) = O(1).\end{aligned}$$

It follows that $c_N'A_N\mathbf{V} = O_p((\frac{N}{h_n})^{\frac{1}{2}})$, by Chebyshev's inequality. □

Appendix C: Proofs for Section 3

Proofs use the following facts: (i) the eigenvalues of a projection matrix are either 0 or 1; (ii) the eigenvalues of a positive definite (p.d.) matrix are strictly positive; (iii) $\gamma_{min}(A)\mathtt{tr}(B) \leq \mathtt{tr}(AB) \leq \gamma_{max}(A)\mathtt{tr}(B)$ for symmetric matrix A and positive semi-definite (p.s.d.) matrix B; (iv) $\gamma_{max}(A+B) \leq \gamma_{max}(A) + \gamma_{max}(B)$ for symmetric matrices A and B; and (v) $\gamma_{max}(AB) \leq \gamma_{max}(A)\gamma_{max}(B)$ for p.s.d. matrices A and B.

Identification Uniqueness: Here, we provide the low-level conditions for the justification of Assumption G. We have $\bar{\sigma}_{v,N}^{*2}(\delta) = \frac{1}{N_1}\eta'\mathbf{A}_N'^{-1}\mathcal{Q}_N'(\delta)\mathcal{Q}_N(\delta)\mathbf{A}_N^{-1}\eta + \frac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbb{Q}_{\mathbb{D}}(\rho)\mathcal{C}_N(\delta)]$, where $\eta = \mathbf{X}\beta_0 + \mathbf{D}\phi_0$, $\mathcal{Q}_N(\delta) = \mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)$, $\mathcal{C}_N(\delta) = \mathbf{C}_N(\delta)(\mathbf{C}_N'\mathbf{C}_N)^{-1}\mathbf{C}_N'(\delta)$ and $\mathbf{C}_N(\delta) = \mathbf{B}_N(\rho)\mathbf{A}_N(\lambda)$. A sufficient condition for Assumption G to hold is either (a) or (b)

holds, where

(a) $\frac{1}{\bar{\sigma}^{*2}_{v,N}(\delta)}\eta'\mathbf{F}'_N\mathbf{B}'_N(\rho)\mathcal{Q}_N(\delta)\mathbf{A}_N^{-1}\eta + \mathtt{tr}[\frac{\sigma^2_{v0}}{\bar{\sigma}^{*2}_{v,N}(\delta)}\mathcal{P}_1(\delta) - \mathcal{P}_2(\delta)] \neq 0$, for $\delta \neq \delta_0$,

(b) $\frac{1}{\bar{\sigma}^{*2}_{v,N}(\delta)}\eta'\mathbf{A}'^{-1}_N\mathcal{Q}'_N(\delta)\mathbf{G}_N(\rho)\mathcal{Q}_N(\delta)\mathbf{A}_N^{-1}\eta + \mathtt{tr}[\frac{\sigma^2_{v0}}{\bar{\sigma}^{*2}_{v,N}(\delta)}\mathcal{P}_3(\rho)\mathcal{C}_N(\delta) - \mathcal{P}_3(\rho)] \neq 0$, for $\delta \neq \delta_0$,

with $\mathcal{P}_1(\delta) = \mathbf{C}'^{-1}_N\mathbf{C}'_N(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)\mathbf{F}_N\mathbf{B}_N^{-1}$, $\mathcal{P}_2(\delta) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)\mathbf{F}_N(\lambda)\mathbf{B}_N^{-1}(\rho)$, and $\mathcal{P}_3(\rho) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{G}_N(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)$. As $\mathcal{Q}_N(\delta_0)\mathbf{A}_N^{-1}\eta = 0$, $\mathcal{C}_N(\delta_0) = I_N$ and $\bar{\sigma}^{*2}_{v,N}(\delta_0) = \sigma^2_{v0}$, the two quantities in (a) and (b) are both 0 at the true parameter values.

Under Assumption H, Lemma B.3 shows that $\|\mathbb{Q}_{\mathbb{D}}(\rho)\|_1$ and $\|\mathbb{Q}_{\mathbb{D}}(\rho)\|_\infty$ are bounded uniformly in $\rho \in \Delta_\rho$, which greatly facilitates the study of the asymptotic properties of M-estimators of spatial parameters. It is in fact not restrictive as it holds for balanced panels. For a balanced panel with a time-invariant and row-normalized spatial weight matrix, we have for all t, $n_t = n$, $D_t = I_n$, $M_t = M$, and $B_t(\rho) = I_n - \rho M \equiv B(\rho)$. As $M \times l_n = l_n$, $J_t(\rho) = I_n - \frac{1}{n}l_n l'_n, t = 2, \ldots, T$. Thus, $B_s(\rho)D_s[\frac{1}{T}\sum_{t=1}^T D'_t B'_t(\rho)J_t(\rho)B_t(\rho)D_t]^{-1}D'_t B'_t(\rho) = (I_n - \frac{T-1}{nT}l_n l'_n)^{-1}$. As $I_n - \frac{T-1}{nT}l_n l'_n$ is strictly diagonally dominant in rows and columns, its inverse is bounded in row and column sum norms (Varah, 1975).

Proof of Theorem 1: By theorem 5.9 of Van der Vaart (1998), we only need to show $\sup_{\delta\in\delta}\frac{1}{N_1}\left\|S^{*c}_N(\delta) - \bar{S}^{*c}_N(\delta)\right\| \xrightarrow{p} 0$ under the assumptions in Theorem 1. From (3.9) and (3.10), the consistency of $\hat{\delta}^*_N$ follows from:

(a) $\inf_{\delta\in\Delta}\bar{\sigma}^{*2}_{v,N}(\delta)$ is bounded away from zero,

(b) $\sup_{\delta\in\Delta}\left|\hat{\sigma}^{*2}_{v,N}(\delta) - \bar{\sigma}^{*2}_{v,N}(\delta)\right| = o_p(1)$,

(c) $\sup_{\delta\in\Delta}\frac{1}{N_1}\left|\mathbf{Y}'\mathbf{W}'\mathbf{B}'_N(\rho)\hat{\mathbf{V}}(\delta) - \mathrm{E}[\mathbf{Y}'\mathbf{W}'\mathbf{B}'_N(\rho)\bar{\mathbf{V}}(\delta)]\right| = o_p(1)$,

(d) $\sup_{\delta\in\Delta}\frac{1}{N_1}\left|\hat{\mathbf{V}}'(\delta)\mathbf{G}_N(\rho)\hat{\mathbf{V}}(\delta) - \mathrm{E}[\bar{\mathbf{V}}'(\delta)\mathbf{G}_N(\rho)\bar{\mathbf{V}}(\delta)]\right| = o_p(1)$.

Proof of (a). Note that $\bar{\beta}^*_N(\delta) = [\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}\mathbb{X}'(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)\mathrm{E}(\mathbf{Y})$ as $\mathbb{X}(\rho) =$

$\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)\mathbf{X}$ and $\mathbb{Q}_{\mathbb{D}}(\rho)$ is idempotent. Thus, $\bar{\mathbf{V}}(\delta) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)\mathbf{Y} - \mathbb{X}(\rho)\bar{\beta}_N^*(\delta) = \mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)\mathbf{Y} + \mathbb{P}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]$. Noting $\mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{P}_{\mathbb{X}}(\rho) = 0$ and $\mathbf{Y} = \mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})$, we have,

$$\begin{aligned}
\bar{\sigma}_{v,N}^{*2}(\delta) &= \tfrac{1}{N_1}\mathrm{E}[\bar{\mathbf{V}}'(\delta)\bar{\mathbf{V}}(\delta)] \\
&= \tfrac{1}{N_1}\mathrm{E}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y}] + \tfrac{1}{N_1}\mathrm{E}\big\{[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]'\mathbf{P}(\delta)[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]\big\} \qquad \text{(C.1)} \\
&= \tfrac{1}{N_1}\mathrm{E}(\mathbf{Y})'\mathbf{Q}(\delta)\mathrm{E}(\mathbf{Y}) + \tfrac{1}{N_1}\mathrm{E}\big\{[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]'[\mathbf{Q}(\delta) + \mathbf{P}(\delta)][\mathbf{Y} - \mathrm{E}(\mathbf{Y})]\big\} \\
&= \tfrac{1}{N_1}\mathrm{E}(\mathbf{Y})'\mathbf{Q}(\delta)\mathrm{E}(\mathbf{Y}) + \tfrac{1}{N_1}\mathrm{E}\big\{[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]'\mathbf{C}_N'(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]\big\} \\
&= \tfrac{1}{N_1}\eta'\mathbf{A}_N'^{-1}\mathbf{Q}(\delta)\mathbf{A}_N^{-1}\eta + \tfrac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbb{Q}_{\mathbb{D}}(\rho)\mathcal{C}_N(\delta)],
\end{aligned}$$

where $\mathbf{Q}(\delta) = \mathbf{C}_N'(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)$ and $\mathbf{P}(\delta) = \mathbf{C}_N'(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbb{P}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)$. The first term can be written in the form of $a'(\delta)a(\delta)$ for an $N \times 1$ vector function of δ, and thus is non-negative, uniformly in $\delta \in \Delta$. For the second term,

$$\begin{aligned}
&\tfrac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbb{Q}_{\mathbb{D}}(\rho)\mathcal{C}_N(\delta)] \geq \tfrac{\sigma_{v0}^2}{N_1}\gamma_{\min}[\mathcal{C}_N(\delta)]\mathtt{tr}[\mathbb{Q}_{\mathbb{D}}(\rho)] = \sigma_{v0}^2\gamma_{\min}[\mathcal{C}_N(\delta)] \\
&\geq \sigma_{v0}^2\gamma_{\max}(\mathbf{A}_N'\mathbf{A}_N)^{-1}\gamma_{\max}(\mathbf{B}_N'\mathbf{B}_N)^{-1}\gamma_{\min}[\mathbf{A}_N'(\lambda)\mathbf{A}_N(\lambda)]\gamma_{\min}[\mathbf{B}_N'(\rho)\mathbf{B}_N(\rho)] > 0,
\end{aligned}$$

uniformly in $\delta \in \Delta$, by Assumption E(iii). It follows that $\inf_{\delta\in\Delta}\bar{\sigma}_{v,N}^{*2}(\delta) > 0$.

Proof of (b). From (3.8), we can write $\hat{\beta}_N^*(\delta) = [\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}\mathbb{X}'(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)\mathbf{Y}$. Then, $\hat{\mathbf{V}}(\delta) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*(\delta)] = \mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)\mathbf{Y}$ and $\hat{\sigma}_{v,N}^{*2}(\delta) = \frac{1}{N_1}\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y}$. From (C.1), $\bar{\sigma}_{v,N}^{*2}(\delta) = \frac{1}{N_1}\mathrm{E}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y}] + \frac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}(\delta)\mathbf{C}_N^{-1}]$. Thus,

$$\hat{\sigma}_{v,N}^{*2}(\delta) - \bar{\sigma}_{v,N}^{*2}(\delta) = \tfrac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})] - \tfrac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}(\delta)\mathbf{C}_N^{-1}].$$

For the second term, $0 \leq \frac{1}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}(\delta)\mathbf{C}_N^{-1}] \leq \frac{1}{N_1}\gamma_{\max}[\mathcal{C}_N(\delta)]\gamma_{\max}^2[\mathbb{Q}_{\mathbb{D}}(\rho)]\mathtt{tr}[\mathbb{P}_{\mathbb{X}}(\rho)] = o(1)$, because $\mathtt{tr}[\mathbb{P}_{\mathbb{X}}(\rho)] = k$, $\gamma_{\max}[\mathbb{Q}_{\mathbb{D}}(\rho)] = 1$ and, by Assumption E(iii),

$$\gamma_{\max}[\mathcal{C}_N(\delta)] \leq \gamma_{\min}(\mathbf{A}_N'\mathbf{A}_N)^{-1}\gamma_{\min}(\mathbf{B}_N'\mathbf{B}_N)^{-1}\gamma_{\max}[\mathbf{A}_N'(\lambda)\mathbf{A}_N(\lambda)]\gamma_{\max}[\mathbf{B}_N'(\rho)\mathbf{B}_N(\rho)] < \infty.$$

Therefore, one has $\sup_{\delta\in\Delta}|\frac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}(\delta)\mathbf{C}_N^{-1}]| = o(1)$. For the first term, we have,

$$
\begin{aligned}
&\tfrac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})] \\
&= \tfrac{1}{N_1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})'\mathbf{A}_N'^{-1}\mathbf{Q}(\delta)\mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V}) - \tfrac{1}{N_1}\mathrm{E}[(\eta + \mathbf{B}_N^{-1}\mathbf{V})'\mathbf{A}_N'^{-1}\mathbf{Q}(\delta)\mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})] \\
&= \tfrac{2}{N_1}\mathbf{V}'\mathbf{C}_N^{-1\prime}\mathbf{Q}(\delta)\mathbf{A}_N^{-1}\eta + \tfrac{1}{N_1}[\mathbf{V}'\mathbf{C}_N^{-1\prime}\mathbf{Q}(\delta)\mathbf{C}_N^{-1}\mathbf{V} - \sigma_{v0}^2\mathtt{tr}(\mathbf{C}_N^{-1\prime}\mathbf{Q}(\delta)\mathbf{C}_N^{-1})].
\end{aligned}
$$

By Assumption E, and Lemmas B.1 and B.3, one shows that $\mathbf{C}_N^{-1\prime}\mathbf{Q}(\delta)\mathbf{A}_N^{-1}$ and $\mathbf{C}_N^{-1\prime}\mathbf{Q}(\delta)\mathbf{C}_N^{-1}$ are bounded in both row and column sum norms, for each $\delta \in \Delta$. Further, the elements of η are uniformly bounded. Thus, the pointwise convergence of the first term follows from Lemma B.6 (v), and the pointwise convergence of the second term follows from Lemma B.6 (iv). Therefore, $\frac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})] \xrightarrow{p} 0$, for each $\delta \in \Delta$.

Next, let δ_1 and δ_2 be in Δ. We have by the mean value theorem (MVT):

$$
\tfrac{1}{N_1}\mathbf{Y}'\mathbf{Q}(\delta_2)\mathbf{Y} - \tfrac{1}{N_1}\mathbf{Y}'\mathbf{Q}(\delta_1)\mathbf{Y} = \tfrac{1}{N_1}\mathbf{Y}'[\tfrac{\partial}{\partial\delta'}\mathbf{Q}(\bar{\delta})]\mathbf{Y}(\delta_2 - \delta_1),
$$

where $\bar{\delta}$ lies between δ_1 and δ_2. It follows that $\frac{1}{N_1}\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y}$ is stochastically equicontinuous if $\sup_{\delta\in\Delta}\frac{1}{N_1}\mathbf{Y}'[\frac{\partial}{\partial\varpi}\mathbf{Q}(\delta)]\mathbf{Y} = O_p(1)$, $\varpi = \lambda, \rho$. We only show $\sup_{\delta\in\Delta}\frac{1}{N_1}\mathbf{Y}'[\frac{\partial}{\partial\rho}\mathbf{Q}(\delta)]\mathbf{Y} = O_p(1)$ as the proof of $\sup_{\delta\in\Delta}\frac{1}{N_1}\mathbf{Y}'[\frac{\partial}{\partial\lambda}\mathbf{Q}(\delta)]\mathbf{Y} = O_p(1)$ is similar and simpler. Note that

$$
\begin{aligned}
\tfrac{\partial}{\partial\rho}\mathbf{Q}(\delta) = &- \mathbf{C}_N'(\delta)\mathbf{G}_N'(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta) + \mathbf{C}_N'(\delta)\dot{\mathbb{Q}}_{\mathbb{D}}(\rho)\mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta) \\
&+ \mathbf{C}_N'(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\dot{\mathbb{Q}}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta) + \mathbf{C}_N'(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbb{Q}_{\mathbb{X}}(\rho)\dot{\mathbb{Q}}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta) \\
&- \mathbf{C}_N'(\delta)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{G}_N(\rho)\mathbf{C}_N(\delta),
\end{aligned}
$$

where $\dot{\mathbb{Q}}_{\mathbb{X}}(\rho) = \frac{\partial}{\partial\rho}\mathbb{Q}_{\mathbb{X}}(\rho)$. Using (A.1), we have after some algebra, $\dot{\mathbb{X}}(\rho) = \frac{\partial}{\partial\rho}\mathbb{X}(\rho) = \mathbb{G}_N(\rho)\mathbb{X}(\rho)$ where $\mathbb{G}_N(\rho) = \mathbb{P}_{\mathbb{D}}(\rho)\mathbf{G}_N'(\rho) - \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{G}_N(\rho)$, which gives

$$
\dot{\mathbb{Q}}_{\mathbb{X}}(\rho) = -\mathbb{P}_{\mathbb{X}}(\rho)\mathbb{G}_N'(\rho)\mathbb{Q}_{\mathbb{X}}(\rho) - \mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{G}_N(\rho)\mathbb{P}_{\mathbb{X}}(\rho). \tag{C.2}
$$

For a conformable vector a and taking use (A.1) and (C.2), we have after some algebra,

$$a'[\frac{\partial}{\partial\rho}\mathbf{Q}(\delta)]a = -2a'\bar{\mathbf{Q}}(\delta)a, \tag{C.3}$$

where $\bar{\mathbf{Q}}(\delta) = \mathcal{Q}'_N(\delta)\mathbf{G}_N(\rho)\mathcal{Q}_N(\delta)$ and $\mathcal{Q}_N(\delta) = \mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)$. Rearranging leads to $\bar{\mathbf{Q}}(\delta) = \mathcal{Q}'_N(\delta)\mathbf{M}\bar{\mathbb{Q}}_{\mathbb{D}}(\rho)\bar{\mathbb{Q}}_{\mathbb{X}}(\rho)\mathbf{A}_N(\lambda)$, where $\bar{\mathbb{Q}}_{\mathbb{D}}(\rho) = I_N - \mathbf{D}[\mathbb{D}'(\rho)\mathbb{D}(\rho)]^{-1}\mathbb{D}'(\rho)\mathbf{B}_N(\rho)$ and $\bar{\mathbb{Q}}_{\mathbb{X}}(\rho) = I_N - \mathbf{X}[\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}\mathbb{X}'(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)$. Following exactly the same way of proving Lemma B.3, we show that $\bar{\mathbb{Q}}_{\mathbb{D}}(\rho)$ and $\bar{\mathbb{Q}}_{\mathbb{X}}(\rho)$ are also bounded in both row and column sum norms, uniformly in $\rho \in \Delta_\rho$. Thus, $\|\bar{\mathbf{Q}}(\delta)\|_1$ and $\|\bar{\mathbf{Q}}(\delta)\|_\infty$ are bounded uniformly in $\delta \in \Delta$. As $\mathbf{Y} = \mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})$, Lemma B.1 and Lemma B.6 imply

$$\frac{1}{N_1}\mathbf{Y}'[\frac{\partial}{\partial\rho}\mathbf{Q}(\delta)]\mathbf{Y} = -\frac{2}{N_1}\mathbf{Y}'\bar{\mathbf{Q}}(\delta)\mathbf{Y} = -\frac{2}{N_1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})'\mathbf{A}_N'^{-1}\bar{\mathbf{Q}}(\delta)\mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})$$
$$= -\frac{2}{N_1}\eta'\mathbf{A}_N'^{-1}\bar{\mathbf{Q}}(\delta)\mathbf{A}_N^{-1}\eta - \frac{4}{N1}\eta'\mathbf{A}_N'^{-1}\bar{\mathbf{Q}}(\delta)\mathbf{C}_N^{-1}\mathbf{V} - \frac{2}{N_1}\mathbf{V}'\mathbf{C}_N'^{-1}\bar{\mathbf{Q}}(\delta)\mathbf{C}_N^{-1}\mathbf{V} = O_p(1),$$

uniformly in $\delta \in \Delta$. Thus, $\sup_{\delta\in\Delta}\frac{1}{N_1}\mathbf{Y}'[\frac{\partial}{\partial\rho}\mathbf{Q}(\delta)]\mathbf{Y} = O_p(1)$. Following a similar analysis, one shows $\sup_{\delta\in\Delta}\frac{1}{N_1}\mathbf{Y}'[\frac{\partial}{\partial\lambda}\mathbf{Q}(\delta)]\mathbf{Y} = O_p(1)$. With the pointwise convergence of $\frac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})]$ to zero for each $\delta \in \Delta$ and the stochastic equicontinuity of $\frac{1}{N_1}\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y}$, the uniform convergence result, $\sup_{\delta\in\Delta}|\frac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})]| = o_p(1)$, follows (Andrews (1992)). Therefore, $\sup_{\delta\in\Delta}|\hat{\sigma}^{*2}_{v,N}(\delta) - \bar{\sigma}^{*2}_{v,N}(\delta)| = o_p(1)$.

Proof of (c). By the expressions of $\hat{\mathbf{V}}(\lambda)$ and $\bar{\mathbf{V}}(\delta)$ given above, we have

$$\frac{1}{N_1}\mathbf{Y}'\mathbf{W}'\mathbf{B}'_N(\rho)\hat{\mathbf{V}}(\delta) - \frac{1}{N_1}\mathrm{E}[\mathbf{Y}'\mathbf{W}'\mathbf{B}'_N(\rho)\bar{\mathbf{V}}(\delta)]$$
$$= \frac{1}{N_1}[\mathbf{Y}'\mathbf{W}'\mathbf{B}'_N(\rho)\mathcal{Q}_N(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{W}'\mathbf{B}'_N(\rho)\mathcal{Q}_N(\delta)\mathbf{Y})] - \frac{\sigma^2_{v0}}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{W}'\mathbf{B}'_N(\rho)\mathcal{P}_N(\delta)\mathbf{C}_N^{-1}],$$

where $\mathcal{P}_N(\delta) = \mathbb{P}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)$. The first term is similar in form to $\frac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})]$ from **(b)** and its uniform convergence is shown in a similar way. Furthermore, by Lemma B.4, it is easy to see that the second term is $o(1)$ uniformly in $\delta \in \Delta$.

Proof of (d). Again, using the expressions of $\bar{\mathbf{V}}(\delta)$ and $\hat{\mathbf{V}}(\delta)$, we have

$$\begin{aligned}&\tfrac{1}{N_1}\hat{\mathbf{V}}'(\delta)\mathbf{G}_N(\rho)\hat{\mathbf{V}}(\delta) - \tfrac{1}{N_1}\mathrm{E}[\bar{\mathbf{V}}'(\delta)\mathbf{G}_N(\rho)\bar{\mathbf{V}}(\delta)]\\ =\ &\tfrac{1}{N_1}[\mathbf{Y}'\bar{\mathbf{Q}}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\bar{\mathbf{Q}}(\delta)\mathbf{Y})] - \tfrac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathcal{P}_N'(\delta)\mathbf{G}_N^{\circ}(\rho)\mathcal{Q}_N(\delta)\mathbf{C}_N^{-1}]\\ &- \tfrac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathcal{P}_N'(\delta)\mathbf{G}_N(\rho)\mathcal{P}_N(\delta)\mathbf{C}_N^{-1}].\end{aligned}$$

Therefore, the uniform convergence of the first term can be shown in a similar way as we do for $\frac{1}{N_1}[\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y} - \mathrm{E}(\mathbf{Y}'\mathbf{Q}(\delta)\mathbf{Y})]$ due to their similar forms. By Lemma B.4, the remaining two terms are easily seen to be $o(1)$, uniformly in $\delta \in \Delta$. □

Proof of Theorem 2: Applying the MVT to each element of $S_N^*(\hat{\theta}_N^*)$, we have

$$0 = \tfrac{1}{\sqrt{N_1}}S_N^*(\hat{\theta}_N^*) = \tfrac{1}{\sqrt{N_1}}S_N^*(\theta_0) + \big[\tfrac{1}{N_1}\tfrac{\partial}{\partial\theta'}S_N^*(\theta)\big|_{\theta=\bar{\theta}_r \text{ in } r\text{th row}}\big]\sqrt{N_1}(\hat{\theta}_N^* - \theta_0), \tag{C.4}$$

where $\{\bar{\theta}_r\}$ are on the line segment between $\hat{\theta}_N^*$ and θ_0. The result follows if

(*a*) $\frac{1}{\sqrt{N_1}}S_N^*(\theta_0) \xrightarrow{D} N[0, \lim_{N\to\infty}\Gamma_N^*(\theta_0)]$,

(*b*) $\frac{1}{N_1}[\frac{\partial}{\partial\theta'}S_N^*(\theta)\big|_{\theta=\bar{\theta}_r \text{ in } r\text{th row}} - \frac{\partial}{\partial\theta'}S_N^*(\theta_0)] = o_p(1)$, and

(*c*) $\frac{1}{N_1}[\frac{\partial}{\partial\theta'}S_N^*(\theta_0) - \mathrm{E}(\frac{\partial}{\partial\theta'}S_N^*(\theta_0))] = o_p(1)$.

Proof of (a). As seen from (3.11), the elements of $S_N^*(\theta_0)$ are linear-quadratic forms in $\mathbf{V}$. Thus, for every non-zero $(k+3)\times 1$ constant vector a, $a'S_N^*(\theta_0)$ is of the form:

$$a'S_N^*(\theta_0) = b_N'\mathbf{V} + \mathbf{V}'\Phi_N\mathbf{V} - \sigma_v^2\mathtt{tr}(\Phi_N),$$

for suitably defined non-stochastic vector b_N and matrix Φ_N. Based on Assumptions A-F, it is easy to verify (by Lemma B.1 and Lemma B.3(i)) that b_N and matrix Φ_N satisfy the conditions of the CLT for LQ form of Kelejian and Prucha (2001), and hence the asymptotic normality of $\frac{1}{\sqrt{N_1}}a'S_N^*(\theta_0)$ follows. By Cramér-Wold device, $\frac{1}{\sqrt{N_1}}S_N^*(\theta_0) \xrightarrow{D} N[0, \lim_{N\to\infty}\Gamma_N^*(\theta_0)]$, where elements of $\Gamma_N^*(\theta_0)$ are given in Appendix A.

Proof of (b). The Hessian matrix $H_N^*(\theta) = \frac{\partial}{\partial\theta'} S_N^*(\theta)$ is given in Appendix A. By Assumptions D and E, and Lemma B.1 and Lemma B.3(i), $\mathbb{R}_N(\rho_0)$, $\mathcal{R}_{1N}(\rho_0)$ and $\mathcal{R}_{2N}(\rho_0)$ are all bounded in row and column sum norms. With these and $\mathbf{Y} = \mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})$, Lemma B.6 leads to $\frac{1}{N_1} H_N^*(\theta_0) = O_p(1)$. Thus, $\frac{1}{N_1} H_N^*(\bar{\theta}) = O_p(1)$ since $\bar{\theta} \xrightarrow{p} \theta_0$ due to $\hat{\theta}_N^* \xrightarrow{p} \theta_0$, where for simplicity, $H_N^*(\bar{\theta})$ is used to denote $\frac{\partial}{\partial\theta'} S_N^*(\theta)\big|_{\theta=\bar{\theta}_r \text{ in } r\text{th row}}$. As $\bar{\sigma}_v^2 \xrightarrow{p} \sigma_{v0}^2$, we have $\bar{\sigma}_v^{-r} = \sigma_{v0}^{-r} + o_p(1)$, for $r = 2, 4, 6$. As σ_v^{-r} appears in $H_N^*(\theta)$ multiplicatively, $\frac{1}{N_1} H_N^*(\bar{\theta}) = \frac{1}{N_1} H_N^*(\bar{\beta}, \bar{\lambda}, \bar{\rho}, \sigma_{v0}^2) + o_p(1)$. Thus, the proof of **(b)** is equivalent to that of

$$\frac{1}{N_1}[H_N^*(\bar{\beta}, \bar{\lambda}, \bar{\rho}, \sigma_{v0}^2) - H_N^*(\theta_0)] \xrightarrow{p} 0,$$

or the proofs of $\frac{1}{N_1}[H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \bar{\rho}, \sigma_{v0}^2) - H_N^{*\mathtt{S}}(\theta_0)] \xrightarrow{p} 0$ and $\frac{1}{N_1}[H_N^{*\mathtt{NS}}(\bar{\delta}) - H_N^{*\mathtt{NS}}(\delta_0)] \xrightarrow{p} 0$, where $H_N^{*\mathtt{S}}$ and $H_N^{*\mathtt{NS}}$ denote, respectively, the stochastic and non-stochastic parts of H_N^*.

For the stochastic part, we see that all the components of $H_N^{*\mathtt{S}}(\beta, \lambda, \rho, \sigma_{v0}^2)$ are linear, bilinear or quadratic in β and λ, but nonlinear in ρ. Hence, with an application of the MVT on $H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \bar{\rho}, \sigma_{v0}^2)$ w.r.t $\bar{\rho}$, we can write $\frac{1}{N_1}[H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \bar{\rho}, \sigma_{v0}^2) - H_N^{*\mathtt{S}}(\theta_0)]$ as

$$\frac{1}{N_1}[\tfrac{\partial}{\partial\rho} H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \dot{\rho}, \sigma_{v0}^2)](\bar{\rho} - \rho_0) + \frac{1}{N_1}[H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \rho_0, \sigma_{v0}^2) - H_N^{*\mathtt{S}}(\theta_0)],$$

where $\dot{\rho}$ lies between $\bar{\rho}$ and ρ_0. Therefore, it suffices to show

$$(i)\ \frac{1}{N_1}\frac{\partial}{\partial\rho} H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \dot{\rho}, \sigma_{v0}^2) = O_p(1) \ \text{ and } \ (ii)\ \frac{1}{N_1}[H_N^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \rho_0, \sigma_{v0}^2) - H_N^{*\mathtt{S}}(\theta_0)] = o_p(1).$$

We do so for the most complicated term, $H_{\rho\lambda}^{*\mathtt{S}}(\theta) = -\frac{1}{\sigma_v^2}\mathbb{Y}'(\rho)\mathbf{G}_N^\circ(\rho)\tilde{\mathbf{V}}(\beta, \delta)$. We have,

$$\frac{1}{N_1}\frac{\partial}{\partial\rho} H_{\rho\lambda}^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \dot{\rho}, \sigma_{v0}^2) = \frac{2}{N_1\sigma_{v0}^2}\mathbb{Y}'(\dot{\rho})\mathcal{R}_{1N}(\dot{\rho})\mathbb{Q}_{\mathbb{D}}(\dot{\rho})\mathbf{B}_N(\dot{\rho})(\mathbf{A}_N(\bar{\lambda})\mathbf{Y} - \mathbf{X}\bar{\beta}),$$

$$\frac{1}{N_1}[H_{\rho\lambda}^{*\mathtt{S}}(\bar{\beta}, \bar{\lambda}, \rho_0, \sigma_{v0}^2) - H_{\rho\lambda}^{*\mathtt{S}}(\theta_0)] = \frac{1}{N_1\sigma_{v0}^2}\mathbb{Y}'\mathbf{G}_N^\circ\mathbb{Y}(\bar{\lambda} - \lambda_0) + \frac{1}{N_1\sigma_{v0}^2}\mathbb{Y}'\mathbf{G}_N^\circ\mathbb{X}(\bar{\beta} - \beta_0).$$

By Lemmas B.1 and B.6, it is easy to show that $\frac{1}{N_1}\mathbb{Y}'\mathbf{G}_N^\circ\mathbb{Y} = O_p(1)$ and $\frac{1}{N_1}\mathbb{Y}'\mathbf{G}_N^\circ\mathbb{X} = O_p(1)$.

Therefore, (ii) holds. To prove (i), we have

$$\mathbb{Y}'(\dot\rho)\mathcal{R}_{1N}(\dot\rho)\mathbb{Q}_{\mathbb{D}}(\dot\rho)\mathbf{B}_N(\dot\rho)(\mathbf{A}_N(\bar\lambda)\mathbf{Y}-\mathbf{X}\bar\beta)$$
$$=(\mathbf{A}_N^{-1}\eta+\mathbf{C}_N^{-1}\mathbf{V})'\mathcal{H}_N(\dot\rho)[\mathbf{A}_N(\bar\lambda)\mathbf{A}_N^{-1}\eta+\mathbf{A}_N(\bar\lambda)\mathbf{C}_N^{-1}\mathbf{V}-\mathbf{X}\bar\beta],$$

where $\mathcal{H}_N(\dot\rho)=\mathbf{W}'\mathbf{B}_N'(\dot\rho)\mathbb{Q}_{\mathbb{D}}(\dot\rho)\mathcal{R}_{1N}(\dot\rho)\mathbb{Q}_{\mathbb{D}}(\dot\rho)\mathbf{B}_N(\dot\rho)$. Lemma B.2 implies $\mathbf{B}_N^{-1}(\dot\rho)$ embedded in $\mathcal{H}_N(\dot\rho)$ is uniformly bounded in both row and column sums since $\dot\rho-\rho_0=o_p(1)$. It follows by Lemma B.6 that the above equation is $O_p(N)$, and then the result (i) follows.

For the non-stochastic part, we illustrate the proof using the most complicated $\lambda\lambda$-term. Noting that the non-stochastic part is nonlinear in both $\bar\lambda$ and $\bar\rho$, we have by the MVT,

$$\tfrac{1}{N_1}[H^{*\mathtt{NS}}_{\lambda\lambda}(\bar\delta)-H^{*\mathtt{NS}}_{\lambda\lambda}(\delta_0)]=-\tfrac{1}{N_1}\mathtt{tr}[\mathbb{Q}_{\mathbb{D}}(\bar\rho)\mathbf{B}_N(\bar\rho)\mathbf{F}_N^2(\bar\lambda)\mathbf{B}_N^{-1}(\bar\rho)-\mathbb{Q}_{\mathbb{D}}\mathbf{B}_N\mathbf{F}_N^2\mathbf{B}_N^{-1}]$$
$$=\;-(\bar\lambda-\lambda_0)\tfrac{1}{N_1}\mathtt{tr}[2\mathbb{Q}_{\mathbb{D}}(\dot\rho)\mathbf{B}_N(\dot\rho)\mathbf{F}_N^3(\dot\lambda)\mathbf{B}_N^{-1}(\dot\rho)]-(\bar\rho-\rho_0)\tfrac{1}{N_1}\mathtt{tr}[\mathbf{F}_N^2(\dot\lambda)\mathbb{R}_N(\dot\rho)],$$

where $\dot\lambda$ lies between $\bar\lambda$ and λ_0 and $\dot\rho$ lies between $\bar\rho$ and ρ_0. Again, by Lemma B.2, we conclude that both $\mathbf{A}_N^{-1}(\dot\lambda)$ and $\mathbf{B}_N^{-1}(\dot\rho)$ are uniformly bounded in both row and column sums. Therefore, the terms inside the trace both have elements that are uniformly bounded. As $\bar\delta-\delta_0=o_p(1)$, we have $\frac{1}{N_1}[H^{*\mathtt{NS}}_{\lambda\lambda}(\bar\delta)-H^{*\mathtt{NS}}_{\lambda\lambda}(\delta_0)]=o_p(1)$.

Proof of (c). Since $\mathbf{Y}=\mathbf{A}_N^{-1}(\eta+\mathbf{B}_N^{-1}\mathbf{V})$, the Hessian matrix at true θ_0 are seen to be linear combinations of terms linear or quadratic in $\mathbf{V}$, and constants. The constant terms are canceled out. Other terms are shown to be $o_p(1)$ based on Lemma B.6. For example,

$$\tfrac{1}{N_1}[H^*_{\rho\rho}(\rho_0)-\mathrm{E}(H^*_{\rho\rho}(\rho_0))]=\tfrac{1}{N_1\sigma_{v0}^2}[\mathbf{V}'\mathbb{Q}_{\mathbb{D}}\mathcal{R}_{1N}\mathbb{Q}_{\mathbb{D}}\mathbf{V}-\mathrm{E}(\mathbf{V}'\mathbb{Q}_{\mathbb{D}}\mathcal{R}_{1N}\mathbb{Q}_{\mathbb{D}}\mathbf{V})]=o_p(1).\qquad\square$$

Corollary C.1. *Under Assumptions A-G, we have,*

$$\Gamma_N^*(\hat\theta_N^*)=\Gamma_N^*(\theta_0)+\mathrm{Bias}^*(\delta_0)+o_p(1),$$

where $\mathrm{Bias}^*(\delta_0)$ *is a matrix having a sole non-zero element* $\frac{1}{N_1}\mathtt{tr}(\mathcal{P}_2'\mathcal{P}_2\mathbb{P}_{\mathbb{D}})$ *at the* λ-λ *entry.*

Proof of Corollary C.1: Note that $\Gamma_N^*(\hat{\theta}_N^*) = \Gamma_N^*(\theta)|_{(\theta=\hat{\theta}_N^*, \phi=\hat{\phi}_N^*, \kappa_3=\hat{\kappa}_{3,N}, \kappa_4=\hat{\kappa}_{4,N})}$. As $\hat{\theta}_N^*$, $\hat{\kappa}_{3,N}$ and $\hat{\kappa}_{4,N}$ are consistent estimators for θ_0, κ_3 and κ_4, plugging these estimators into $\Gamma_N^*(\theta)$ will not bring additional bias to the estimation of $\Gamma_N^*(\theta_0)$. However, due to incidental parameters problem, the $\hat{\mu}_N^*$ component of $\hat{\phi}_N^*$ is not consistent for the estimation of μ_0 when T is fixed. To estimate the bias caused by replacing ϕ_0 by $\hat{\phi}_N^*$, recall (3.3),

$$\hat{\phi}_N(\beta, \delta) = [\mathbb{D}'(\rho)\mathbb{D}(\rho)]^{-1}\mathbb{D}'(\rho)\mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\beta].$$

Thus, the unconstrained estimate of ϕ_0 is just $\hat{\phi}_N^* = \hat{\phi}_N(\hat{\beta}_N^*, \hat{\delta}_N^*)$. Note $\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^* = \mathbf{A}_N\mathbf{Y} - \mathbf{X}\beta_0 - \mathbf{W}\mathbf{Y}(\hat{\lambda}_N^* - \lambda_0) - \mathbf{X}(\hat{\beta}_N^* - \beta_0)$. Applying the MVT on each row of $\mathbf{D}\hat{\phi}_N^*$ with respect to the $\hat{\rho}_N^*$-element, we have,

$$\begin{aligned}
\mathbf{D}\hat{\phi}_N^* &= \mathbf{D}[\mathbb{D}'(\hat{\rho}_N^*)\mathbb{D}(\hat{\rho}_N^*)]^{-1}\mathbb{D}'(\hat{\rho}_N^*)\mathbf{B}_N(\hat{\rho}_N^*)[\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*] \\
&= \mathbf{B}_N^{-1}(\hat{\rho}_N^*)\mathbb{P}_{\mathbb{D}}(\hat{\rho}_N^*)\mathbf{B}_N(\hat{\rho}_N^*)[\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*] \\
&= [\mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{B}_N - \mathbb{R}_N(\bar{\rho})(\hat{\rho}_N^* - \rho_0)][\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*] \\
&= \mathbf{D}\phi_0 + \mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{V} - \mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{B}_N[\mathbf{W}\mathbf{Y}(\hat{\lambda}_N^* - \lambda_0) + \mathbf{X}(\hat{\beta}_N^* - \beta_0)] \\
&\quad - \mathbb{R}_N(\bar{\rho})[\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*](\hat{\rho}_N^* - \rho_0),
\end{aligned} \tag{C.5}$$

where $\bar{\rho}$ lies between $\hat{\rho}_N^*$ and ρ_0 and changes over the rows of $\mathbb{R}_N(\bar{\rho})$, and $\mathbb{R}_N(\rho)$ is given in the appendix of the main text. From its expression, $\Gamma_N^*(\theta)$ is seen to have components linear or quadratic in $\mathbf{D}\phi$. Let d_N be a non-stochastic N-vector with elements being of uniform order $O(1)$ or $O(h_n^{-1})$. Using (C.5), the terms of $\Gamma_N^*(\hat{\theta}_N^*)$ linear in $\mathbf{D}\hat{\phi}_N^*$ are represented as

$$\begin{aligned}
\tfrac{1}{N_1}d_N'\mathbf{D}\hat{\phi}_N^* &= \tfrac{1}{N_1}d_N'\mathbf{D}\phi_0 + \tfrac{1}{N_1}d_N'\mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{V} - \tfrac{1}{N_1}d_N'\mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{B}_N[\mathbf{W}\mathbf{Y}(\hat{\lambda}_N^* - \lambda_0) + \mathbf{X}(\hat{\beta}_N^* - \beta_0)] \\
&\quad + \tfrac{1}{N_1}d_N'\mathbb{R}_N(\bar{\rho})[\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*](\hat{\rho}_N^* - \rho_0) = \tfrac{1}{N_1}d_N'\mathbf{D}\phi_0 + o_p(1),
\end{aligned}$$

where the last equation holds because of the consistency of $\hat{\theta}_N^*$ and Lemma B.6, using $\mathbf{Y} =$

$\mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})$. Hence, we can conclude that the terms of $\Gamma_N^*(\theta_0)$ linear in ϕ_0 can be consistently estimated by simply replacing ϕ_0 with $\hat{\phi}_N^*$.

The only term that is quadratic in ϕ_0 is contained in $\Gamma_{\lambda\lambda}^*(\theta_0)$, $\frac{1}{N_1\sigma_{v0}^2}\phi_0'\mathbb{D}'\mathcal{P}_2'\mathcal{P}_2\mathbb{D}\phi_0$. The plug-in estimator estimates this term by $\frac{1}{N_1\hat{\sigma}_{v,N}^{*2}}\hat{\phi}_N^{*\prime}\mathbb{D}'(\hat{\rho}_N^*)\mathcal{P}_2'(\hat{\delta}_N^*)\mathcal{P}_2(\hat{\delta}_N^*)\mathbb{D}(\hat{\rho}_N^*)\hat{\phi}_N^*$. Using (C.5), $\hat{\theta}_N^* - \theta_0 = o_p(1)$ and Lemma B.6, we show that this estimator is biased/inconsistent:

$$
\begin{aligned}
&\frac{1}{N_1\hat{\sigma}_{v,N}^{*2}}\hat{\phi}_N^{*\prime}\mathbb{D}'(\hat{\rho}_N^*)\mathcal{P}_2'(\hat{\delta}_N^*)\mathcal{P}_2(\hat{\delta}_N^*)\mathbb{D}(\hat{\rho}_N^*)\hat{\phi}_N^* \\
&= \frac{1}{N_1\hat{\sigma}_{v,N}^{*2}}\phi_0'\mathbb{D}'(\hat{\rho}_N^*)\mathcal{P}_2'(\hat{\delta}_N^*)\mathcal{P}_2(\hat{\delta}_N^*)\mathbb{D}(\hat{\rho}_N^*)\phi_0 \\
&\quad + \frac{1}{N_1\hat{\sigma}_{v,N}^{*2}}\mathbf{V}'\mathbb{P}_{\mathbb{D}}\mathbf{B}_N^{-1\prime}\mathbf{B}_N'(\hat{\rho}_N^*)\mathcal{P}_2'(\hat{\delta}_N^*)\mathcal{P}_2(\hat{\delta}_N^*)\mathbf{B}_N(\hat{\rho}_N^*)\mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{V} + o_p(1) \\
&= \frac{1}{N_1\sigma_{v0}^2}\phi_0'\mathbb{D}'\mathcal{P}_2'\mathcal{P}_2\mathbb{D}\phi_0 + \frac{1}{N_1\sigma_{v0}^2}\mathbf{V}'\mathbb{P}_{\mathbb{D}}\mathcal{P}_2'\mathcal{P}_2\mathbb{P}_{\mathbb{D}}\mathbf{V} + o_p(1) \\
&= \frac{1}{N_1\sigma_{v0}^2}\phi_0'\mathbb{D}'\mathcal{P}_2'\mathcal{P}_2\mathbb{D}\phi_0 + \frac{1}{N_1}\mathtt{tr}[\mathcal{P}_2'\mathcal{P}_2\mathbb{P}_{\mathbb{D}}] + o_p(1).
\end{aligned}
$$

We see that the bias term, $\frac{1}{N_1}\mathtt{tr}[\mathcal{P}_2'\mathcal{P}_2\mathbb{P}_{\mathbb{D}}]$, involves only the common parameters that can be consistently estimated. Thus, a bias correction can easily be made. Define

$$
\text{Bias}_{\lambda\lambda}^*(\delta) = \frac{1}{N_1}\mathtt{tr}[\mathcal{P}_2'(\delta)\mathcal{P}_2(\delta)\mathbb{P}_{\mathbb{D}}(\rho)]. \tag{C.6}
$$

This gives the bias matrix of $\Gamma_N^*(\hat{\theta}_N^*)$, which is a matrix of the same dimension as $\Gamma_N^*(\hat{\theta}_N^*)$, and has the sole non-zero element $\text{Bias}_{\lambda\lambda}^*(\delta_0)$ corresponding to the $\Gamma_{\lambda\lambda}^*(\hat{\theta}_N^*)$ component. □

Corollary C.2. *Under Assumptions A-G, we have, as $N \to \infty$,*

(i) $\hat{\kappa}_{3,N} \xrightarrow{p} \kappa_{3,0}$ *and* $\hat{\kappa}_{4,N} \xrightarrow{p} \kappa_{4,0}$; *(ii)* $\widehat{\Sigma}_N^* - \Sigma_N^*(\theta_0) \xrightarrow{p} 0$ *and* $\widehat{\Gamma}_N^* - \Gamma_N^*(\theta_0) \xrightarrow{p} 0$;

and therefore $\widehat{\Sigma}_N^{*-1}\widehat{\Gamma}_N^*\widehat{\Sigma}_N^{*\prime-1} - \Sigma_N^{*-1}(\theta_0)\Gamma_N^*(\theta_0)\Sigma_N^{*\prime-1}(\theta_0) \xrightarrow{p} 0$.

Proof of Corollary C.2.

Proof of (i). Note: $\mathbf{V} = \mathbf{B}_N(\mathbf{A}_N\mathbf{Y} - \eta)$, $\tilde{\mathbf{V}} = \mathbb{Q}_{\mathbb{D}}\mathbf{V}$ and $\hat{\mathbf{V}} = \mathbb{Q}_{\mathbb{D}}(\hat{\rho}_N^*)\mathbf{B}_N(\hat{\rho}_N^*)[\mathbf{A}_N(\hat{\lambda}_N^*)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^*]$ with respective elements $\{v_j\}, \{\tilde{v}_j\}$ and $\{\hat{v}_j\}$, and $\mathbb{Q}_{\mathbb{D}}$ has elements $\{q_{jh}\}$, $j, h = 1, \ldots, N$, where j and h are the combined indices for $i = 1, \ldots, n_t$ and $t = 1, \ldots, T$.

Consistency of $\hat{\kappa}_{3,N}$. As $\hat{\sigma}^*_{v,N} - \sigma_{v0} = o_p(1)$ and $\hat{\rho}^*_N - \rho_0 = o_p(1)$, the denominators of $\hat{\kappa}_{3,N}$ and κ_3 agree asymptotically. Thus, $\hat{\kappa}_{3,N}$ is consistent if $\frac{1}{N}\sum_{j=1}^N[\hat{v}_j^3 - \mathrm{E}(\tilde{v}_j^3)] \xrightarrow{p} 0$, or

(a) $\frac{1}{N}\sum_{j=1}^N[\tilde{v}_j^3 - \mathrm{E}(\tilde{v}_j^3)] \xrightarrow{p} 0$ and (b) $\frac{1}{N}\sum_{j=1}^N(\hat{v}_j^3 - \tilde{v}_j^3) \xrightarrow{p} 0$.

To prove (a), noting that $\tilde{v}_j = \sum_{h=1}^N q_{jh}v_h$, we have,

$$\frac{1}{N}\sum_{j=1}^N[\tilde{v}_j^3 - \mathrm{E}(\tilde{v}_j^3)] = \frac{1}{N}\sum_{j=1}^N\sum_{h=1}^N q_{jh}^3[v_h^3 - \mathrm{E}(v_h^3)] + \frac{3}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N q_{jl}^2 q_{jm}v_l^2 v_m$$
$$+\frac{6}{N}\sum_{j=1}^N\sum_{m=1}^N\sum_{\substack{l\neq m\\ l=1}}^N\sum_{\substack{h\neq m,l\\ h=1}}^N q_{jm}q_{jl}q_{jh}v_m v_l v_h \equiv K_1 + K_2 + K_3.$$

First, consider K_1 term. By Lemma B.3, $\mathbb{Q}_{\mathbb{D}}$ is uniformly bounded in both row and column sums. This implies that the elements of $\mathbb{Q}_{\mathbb{D}}$ are uniformly bounded. Therefore, there exists a constant $\bar{q}$ such that $|q_{jh}| \leq \bar{q}$ for all j and h. Given these, we have $\sum_{j=1}^N q_{jh}^3 \leq \sum_{j=1}^N |q_{jh}|^3 \leq \bar{q}^2\sum_{j=1}^N |q_{jh}| < \infty$. Also note $\{v_i\}$ are iid by Assumption A. Thus, Khinchine's weak law of large number (WLLN) (Feller, 1967, pp. 243-244) implies that K_1 converges to zero in probability as the sample size increases.

For the other two terms, we have by switching the order of summations when needed,

$$K_2 = \frac{3}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N q_{jl}^2 q_{jm}(v_l^2 - \sigma_v^2)v_m + \frac{3}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N q_{jl}^2 q_{jm}\sigma_v^2 v_m$$
$$= \frac{3}{N}\sum_{m=1}^N(v_m^2 - \sigma_v^2)(\sum_{j=1}^N\sum_{l=1}^{m-1} q_{jm}^2 q_{jl}v_l) + \frac{3}{N}\sum_{m=1}^N v_m[\sum_{j=1}^N\sum_{l=1}^{m-1} q_{jl}^2 q_{jm}(v_l^2 - \sigma_v^2)]$$
$$+\frac{3}{N}\sum_{m=1}^N\sum_{j=1}^N\sum_{\substack{l\neq m\\ l=1}}^N q_{jl}^2 q_{jm}\sigma_v^2 v_m \equiv \frac{1}{N}\sum_{m=1}^N(g_{1,m} + g_{2,m} + g_{3,m}), \text{ and}$$

$$K_3 = \frac{18}{N}\sum_{m=1}^N v_m(\sum_{j=1}^N\sum_{l=1}^{m-1}\sum_{\substack{h\neq l\\ h=1}}^{m-1} q_{jm}q_{jl}q_{jh}v_l v_h) \equiv \frac{1}{N}\sum_{m=1}^N g_{4,m},$$

where $g_{1,m} = 3(v_m^2 - \sigma_v^2)\sum_{j=1}^N\sum_{l=1}^{m-1} q_{jm}^2 q_{jl}v_l$, $g_{2,m} = 3v_m\sum_{j=1}^N\sum_{l=1}^{m-1} q_{jl}^2 q_{jm}(v_l^2 - \sigma_v^2)$, $g_{3,m} = 3\sum_{j=1}^N\sum_{\substack{l\neq m\\ l=1}}^N q_{jl}^2 q_{jm}\sigma_v^2 v_m$, and $g_{4,m} = v_m\sum_{j=1}^N\sum_{l=1}^{m-1}\sum_{\substack{h\neq l\\ h=1}}^{m-1} q_{jm}q_{jl}q_{jh}v_l v_h$.

Let $\{\mathcal{G}_m\}$ be the increasing sequence of σ-fields generated by $(v_1, \cdots, v_j, j = 1, \cdots, m)$, $m = 1, \cdots, N$. Then, $\mathrm{E}[(g_{1,m}, g_{2,m}, g_{3,m}, g_{4,m})|\mathcal{G}_{m-1}] = 0$; hence, $\{(g_{1,m}, g_{2,m}, g_{3,m}, g_{4,m})', \mathcal{G}_m\}$ form a vector martingale difference (M.D.) sequence. As $\mathbb{Q}_{\mathbb{D}}$ is bounded in row and column sum norms, by Assumption A, it is easy to see that $\mathrm{E}|g_{s,m}|^{1+\epsilon} < \infty$, for $s = 1, 2, 3, 4$ and $\epsilon > 0$. Hence, $\{g_{1,m}\}$, $\{g_{2,m}\}$, $\{g_{3,m}\}$ and $\{g_{4,m}\}$ are uniformly integrable, and the WLLN of Davidson (1994, Theorem 19.7) applies to give $K_2 \xrightarrow{p} 0$ and $K_3 \xrightarrow{p} 0$.

To prove (b), using the notation $\tilde{\mathbf{V}}(\xi) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\beta]$ in (3.4) where $\xi = (\beta', \delta')'$, we have $\tilde{\mathbf{V}} = \tilde{\mathbf{V}}(\xi_0)$ and $\hat{\mathbf{V}} = \tilde{\mathbf{V}}(\hat{\xi}_N^*)$. Let $\mathbf{S}(\xi) = \frac{\partial}{\partial \xi'}\tilde{\mathbf{V}}(\xi)$, and we have

$$\mathbf{S}(\xi) = \{-\mathbb{X}(\rho),\ \ -\mathbb{Y}(\rho),\ \ [\dot{\mathbb{Q}}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho) - \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{M}][\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\beta]\},$$

where the expression of $\dot{\mathbb{Q}}_{\mathbb{D}}(\rho)$ is in (A.1) and $\mathbb{Y}(\rho) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)\mathbf{W}\mathbf{Y}$. Let $s_j'(\xi)$ be the jth row of $\mathbf{S}(\xi)$. We have by the MVT, for each $j = 1, 2, \ldots, N$,

$$\hat{v}_j \equiv \tilde{v}_j(\hat{\xi}_N^*) = \tilde{v}_j(\xi_0) + s_j'(\bar{\xi})(\hat{\xi}_N^* - \xi_0) = \tilde{v}_j + \psi_j'(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|), \tag{C.7}$$

where $\bar{\xi}$ lies between $\hat{\xi}_N^*$ and ξ_0, and $\psi_j' = \mathrm{plim}_{N\to\infty} s_j'(\bar{\xi})$, which is easily shown to be $O_p(1)$ as follows. Consider the first k (the number of regressors) elements of ψ_j' first. They are the limits of the jth row of $-\mathbb{X}(\bar{\rho})$, which are just the jth row of $-\mathbb{X}$ because $\bar{\rho} \xrightarrow{p} \rho_0$, implied by $\hat{\rho}_N^* - \rho_0 = o_p(1)$. Hence, we conclude that the first k elements of ψ_j' are $O(1)$, for each $j = 1, 2, \ldots, N$. For the remaining two elements in each ψ_j', they are the limits of elements from the last two columns of $\mathbf{S}(\bar{\xi})$. It is easy to see the limits of the last two columns of $\mathbf{S}(\bar{\xi})$ are just $-\mathbb{Y}$ and $[\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}][\mathbf{A}_N\mathbf{Y} - \mathbf{X}\beta_0]$. Using $\mathbf{Y} = \mathbf{A}_N^{-1}\eta + \mathbf{C}_N^{-1}\mathbf{V}$, we have $-\mathbb{Y} = \mathcal{P}_2\mathbf{B}_N\eta + \mathcal{P}_2\mathbf{V}$ and $[\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}][\mathbf{A}_N\mathbf{Y} - \mathbf{X}\beta_0] = [\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}]\mathbf{D}\phi_0 + [\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}]\mathbf{B}_N^{-1}\mathbf{V}$. By Lemma B.1, we have the elements of $\mathcal{P}_2\mathbf{B}_N\eta$ and $[\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}]\mathbf{D}\phi_0$ are uniformly bounded, and $\mathcal{P}_2$ and $[\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}]\mathbf{B}_N^{-1}$ are uniformly bounded in both row and column sum norms. Hence, it is easy to see each element of $-\mathbb{Y}$ and $[\dot{\mathbb{Q}}_{\mathbb{D}}\mathbf{B}_N - \mathbb{Q}_{\mathbb{D}}\mathbf{M}][\mathbf{A}_N\mathbf{Y} - \mathbf{X}\beta_0]$ are $O_p(1)$, i.e., the last two elements in ψ_j' are also $O_p(1)$, for each $j = 1, 2, \ldots, N$.

As $\tilde{v}_j = O_p(1)$, $\psi_j' = O_p(1)$ and $\hat{\xi}_N^* - \xi_0 = O_p(\frac{1}{\sqrt{N_1}})$, we have by (C.7), $\hat{v}_j^3 = \tilde{v}_j^3 +$

$3\tilde{v}_j^2\psi_j'(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|)$. It follows that

$$\frac{1}{N}\sum_{j=1}^N(\hat{v}_j^3 - \tilde{v}_j^3) = \frac{3}{N}\sum_{j=1}^N \tilde{v}_j^2\psi_j'(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|)$$
$$= \frac{3\sigma_v^2}{N}\sum_{j=1}^N(\sum_{k=1}^N q_{jk}^2\psi_j')(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|) = o_p(1),$$

as $\frac{1}{N}\sum_{j=1}^N(\sum_{k=1}^N q_{jk}^2\psi_j') = (\sum_{k=1}^N q_{jk}^2)\frac{1}{N}(\sum_{j=1}^N \psi_j') = O(1)$.

Consistency of $\hat{\kappa}_{4,N}$. As $\hat{\sigma}_{v,N}^* - \sigma_{v0} = o_p(1)$ and $\hat{\rho}_N^* - \rho_0 = o_p(1)$, the result follows if $\frac{1}{N}\sum_{j=1}^N[\hat{v}_j^4 - \mathrm{E}(\tilde{v}_j^4)] \xrightarrow{p} 0$. This shows that

$$(c)\ \frac{1}{N}\sum_{j=1}^N[\tilde{v}_j^4 - \mathrm{E}(\tilde{v}_j^4)] \xrightarrow{p} 0 \ \text{ and } \ (d)\ \frac{1}{N}\sum_{j=1}^N(\hat{v}_j^4 - \tilde{v}_j^4) \xrightarrow{p} 0.$$

To prove (c), we have

$$\frac{1}{N}\sum_{j=1}^N \tilde{v}_j^4 - \frac{1}{N}\sum_{j=1}^N \mathrm{E}(\tilde{v}_j^4)$$
$$= \frac{1}{N}\sum_{j=1}^N\sum_{h=1}^N q_{jh}^4[v_h^4 - \mathrm{E}(v_h^4)] + \frac{3}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N q_{jl}^2 q_{jm}^2(v_l^2 v_m^2 - \sigma_v^4)$$
$$+ \frac{4}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N q_{jl}^3 q_{jm} v_l^3 v_m + \frac{6}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N\sum_{\substack{h\neq m,l\\ h=1}}^N q_{jl}^2 q_{jm} q_{jh} v_l^2 v_m v_h$$
$$+ \frac{1}{N}\sum_{j=1}^N\sum_{l=1}^N\sum_{\substack{m\neq l\\ m=1}}^N\sum_{\substack{h\neq m,l\\ h=1}}^N\sum_{\substack{p\neq m,l,h\\ p=1}}^N q_{jl}q_{jm}q_{jh}q_{jp}v_l v_m v_h v_p \equiv \sum_{r=1}^5 R_r.$$

By using WLLN of Davidson (1994, Theorem 19.7) for M.D. arrays as in the proof of (a), we have $R_r = o_p(1)$ for $r = 1, 3, 4, 5$. For R_2, we have

$$R_2 = \frac{6}{N}\sum_{l=1}^N(v_l^2 - \sigma_v^2)[\sum_{j=1}^N\sum_{m=1}^{l-1} q_{jl}^2 q_{jm}^2(v_m^2 - \sigma_v^2)]$$
$$+ \frac{6}{N}\sum_{l=1}^N[\sum_{j=1}^N\sum_{\substack{m\neq l\\ m=1}}^N q_{jl}^2 q_{jm}^2\sigma_v^2(v_l^2 - \sigma_v^2)] \equiv \frac{6}{N}\sum_{l=1}^N(f_{1,l} + f_{2,l}),$$

noting that $v_l^2 v_m^2 - \sigma_v^4 = (v_l^2 - \sigma_v^2)(v_m^2 - \sigma_v^2) + \sigma_v^2(v_m^2 - \sigma_v^2) + \sigma_v^2(v_l^2 - \sigma_v^2)$. Since $\mathrm{E}[f_{1,l}|\mathcal{G}_{l-1}] = 0$ and $\{f_{2,l}\}$ are independent, both $\{f_l\}$ and $\{f_{2,l}\}$ form M.D. sequences. It is easy to see that $\mathrm{E}|f_{s,l}|^{1+\epsilon} < \infty$, for $s = 1, 2$ and $\epsilon > 0$, so that $\{f_{1,l}\}$ and $\{f_{2,l}\}$ are uniformly integrable. Therefore, the WLLN of Davidson (1994, Theorem 19.7) implies that $\frac{6}{N}\sum_{l=1}^N f_{1,l} = o_p(1)$ and $\frac{6}{N}\sum_{l=1}^N f_{2,l} = o_p(1)$.

To prove (d), $\hat{v}_j^4 = \tilde{v}_j^4 + 4\tilde{v}_j^3\psi_j'(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|)$ by (C.7). It follows that

$$\begin{aligned}\tfrac{1}{N}\textstyle\sum_{j=1}^N(\hat{v}_j^4 - \tilde{v}_j^4) &= \tfrac{4}{N}\textstyle\sum_{j=1}^N \tilde{v}_j^3\psi_j'(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|)\\ &= \tfrac{4\sigma_v^3\kappa_3}{N}\textstyle\sum_{j=1}^N(\sum_{k=1}^N q_{jk}^3\psi_j')(\hat{\xi}_N^* - \xi_0) + o_p(\|\hat{\xi}_N^* - \xi_0\|) = o_p(1).\end{aligned}$$

Proof of (ii). The consistency of $\widehat{\Sigma}_N^*$ to $\Sigma_N^*(\theta_0)$ can be shown similarly as what we do in the proof of Theorem 2 for results (b) and (c). For $\widehat{\Gamma}_N^* - \Gamma_N^*(\theta_0) \xrightarrow{p} 0$, we only need to show that $\text{Bias}^*(\hat{\delta}_N^*) - \text{Bias}^*(\delta_0) = o_p(1)$, based on Corollary C.1. That is to show

$$\tfrac{1}{N_1}\{\texttt{tr}[\mathcal{P}_2'(\hat{\delta}_N^*)\mathcal{P}_2(\hat{\delta}_N^*)\mathbb{P}_{\mathbb{D}}(\hat{\rho}_N^*)] - \texttt{tr}(\mathcal{P}_2'\mathcal{P}_2\mathbb{P}_{\mathbb{D}})\} = o_p(1),$$

which can be proved as that for $\frac{1}{N_1}[H_{\lambda\lambda}^{*\texttt{NS}}(\bar{\delta}) - H_{\lambda\lambda}^{*\texttt{NS}}(\delta_0)]$ in the proof of Theorem 2 (b). □

Appendix D: Proofs for Section 4

Identification Uniqueness: Let $\bar{S}_N^\diamond(\beta,\delta) = \text{E}[S_N^\diamond(\beta,\delta)]$ be the population robust AQS functions. Then, the β-component of $\bar{S}_N^\diamond(\beta,\delta) = 0$ is solved at $\bar{\beta}_N^\diamond(\delta) = \bar{\beta}_N^*(\delta)$ given below (3.10). Upon substitution, we obtain the population counterpart of $S_N^{\diamond c}(\delta)$:

$$\bar{S}_N^{\diamond c}(\delta) = \begin{cases}\text{E}[\mathbf{Y}'\mathbf{C}_N'(\delta)[\bar{\mathbf{F}}_N'(\delta) - \bar{\mathbb{F}}_N'(\delta)]\bar{\mathbf{V}}(\delta)],\\ \text{E}\big\{[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\bar{\beta}_N^\diamond(\delta)]'\mathbf{B}_N'(\rho)[\bar{\mathbf{G}}_N(\rho) - \bar{\mathbb{G}}_N(\rho)]\bar{\mathbf{V}}(\delta)\big\},\end{cases} \tag{D.1}$$

where $\bar{\mathbf{V}}(\delta) = \tilde{\mathbf{V}}(\bar{\beta}_N^\diamond(\delta),\delta)$. As $S_N^{\diamond c}(\hat{\delta}_N^\diamond)$ and $\bar{S}_N^{\diamond c}(\delta_0)$ are both zero, by Theorem 5.9 of Van der Vaart (1998) $\hat{\delta}_N^\diamond$ will be consistent if $\sup_{\delta\in\Delta}\frac{1}{N_1}\left\|S_N^{\diamond c}(\delta) - \bar{S}_N^{\diamond c}(\delta)\right\| \xrightarrow{p} 0$ and Assumption G′ holds. Next, we show Assumption G′ holds if either of the following conditions holds:

(a) $\eta'\mathbf{A}_N'^{-1}\mathbf{C}_N'(\delta)[\bar{\mathbf{F}}_N'(\delta) - \bar{\mathbb{F}}_N'(\delta)]\mathcal{Q}_N(\delta)\mathbf{A}_N^{-1}\eta + \texttt{tr}[\mathbb{Q}_{\mathbb{D}}(\rho)\mathcal{C}_N^h(\delta)\big(\bar{\mathbf{F}}_N'(\delta) - \bar{\mathbb{F}}_N'(\delta)\big)] \neq 0$; or

(b) $\eta'\mathbf{A}_N'^{-1}\mathbf{C}_N'(\delta)\mathbb{M}_N'(\rho)\left[\bar{\mathbf{G}}_N(\rho) - \bar{\mathbb{G}}_N(\rho)\right]\mathcal{Q}_N(\delta)\mathbf{A}_N^{-1}\eta + \texttt{tr}[\mathbb{Q}_{\mathbb{D}}(\rho)\mathcal{C}_N^h(\delta)\big(\bar{\mathbf{G}}_N(\rho) - \bar{\mathbb{G}}_N(\rho)\big)] \neq 0$,

for $\delta \neq \delta_0$, where $\mathcal{C}_N^h(\delta) = \mathbf{C}_N(\delta)\mathbf{C}_N^{-1}\mathbf{H}\mathbf{C}_N^{-1\prime}\mathbf{C}_N'(\delta)$ and $\mathbb{M}_N(\rho) = \mathbf{I}_N - \mathbf{B}_N(\rho)\mathbf{X}[\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}\mathbb{X}'(\rho)$. As $\mathcal{C}_N^h(\delta_0) = \mathbf{H}$ and $\mathcal{Q}_N(\delta_0)\mathbf{A}_N^{-1}\eta = 0$, the two quantities in (a) and (b) are 0 at δ_0.

Validity of the assumptions on $\Pi_N(\rho)$ in Lemma D.1 under a balanced panel:

Following the first part in Appendix C, we have $\mathbb{Q}_{\mathbb{D}}(\rho) = (I_T - \frac{l_T l_T'}{T}) \otimes (I_n - \frac{l_n l_n'}{n})$, where $\otimes$ denotes the Kronecker product. Thus, $[\mathbb{Q}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho)]^{-1}$ exists if $T > 2$ by Schur product theorem. Further, $|(\mathbb{Q}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho))_{ii}| - \sum_{j \neq i} |(\mathbb{Q}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho))_{ij}| = \frac{(n-1)(T-1)[(n-2)(T-2)-2]}{n^2 T^2} > c > 0, \forall i,\ T > 2$. As $\mathbb{Q}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho)$ is symmetric, we conclude it is strictly diagonally dominant in both rows and columns. Hence, Theorem 1 and Corollary 1 of Varah (1975) imply that $\|\Pi_N(\rho)\|_1$ and $\|\Pi_N(\rho)\|_\infty$ are both bounded.

Proof of Theorem 3. Consistency of $\hat{\beta}_N^{\diamond}$ follows that of $\hat{\delta}_N^{\diamond}$ under Assumptions C and E. Thus, we only need to prove the consistency of $\hat{\delta}_N^{\diamond}$. By theorem 5.9 of Van der Vaart (1998), $\hat{\delta}_N^{\diamond}$ will be consistent for δ_0 if $\sup_{\delta \in \Delta} \frac{1}{N_1} \left\| S_N^{\diamond c}(\delta) - \bar{S}_N^{\diamond c}(\delta) \right\| \xrightarrow{p} 0$.

Let $\mathbb{L}_\lambda(\delta) = \mathbb{Q}_{\mathbb{D}}(\rho)[\bar{\mathbf{F}}_N(\delta) - \bar{\mathbb{F}}_N(\delta)]$, $\mathbb{L}_\rho(\rho) = \mathbb{Q}_{\mathbb{D}}(\rho)[\bar{\mathbf{G}}_N'(\rho) - \bar{\mathbb{G}}_N'(\rho)]$ and $\mathbb{N}_N(\rho) = I_N - \mathbb{M}_N(\rho)$. Note that $\mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\hat{\beta}_N^{\diamond}(\delta)] = \mathbb{M}_N(\rho)\mathbf{C}_N(\delta)\mathbf{Y}$ and $\mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\bar{\beta}_N^{\diamond}(\delta)] = \mathbb{M}_N(\rho)\mathbf{C}_N(\delta)\mathbf{Y} + \mathbb{N}_N(\rho)\mathbf{C}_N(\delta)[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]$. Recall $\hat{\mathbf{V}}(\delta) = \mathcal{Q}_N(\delta)\mathbf{Y}$ and $\bar{\mathbf{V}}(\delta) = \mathcal{Q}_N(\delta)\mathbf{Y} + \mathcal{P}_N(\delta)[\mathbf{Y} - \mathrm{E}(\mathbf{Y})]$. With Assumption G′, consistency of $\hat{\delta}_N^{\diamond}$ follows if:

(i) $\sup_{\delta \in \Delta} \frac{1}{N_1} \left| \mathbf{Y}'\mathbf{Q}_r^h(\delta)\mathbf{Y} - \mathrm{E}[\mathbf{Y}'\mathbf{Q}_r^h(\delta)\mathbf{Y}] \right| = o_p(1)$, for $r = 1, 2$;

(ii) $\sup_{\delta \in \Delta} \frac{\sigma_{v0}^2}{N_1} \mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}_s^h(\delta)\mathbf{C}_N^{-1}] = o(1)$, for $s = 1, 2, 3$;

where $\mathbf{Q}_1^h(\delta) = \mathbf{C}_N'(\delta)\mathbb{L}_\lambda'(\delta)\mathcal{Q}_N(\delta)$, $\mathbf{Q}_2^h(\delta) = \mathbf{C}_N'(\delta)\mathbb{M}_N'(\rho)\mathbb{L}_\rho'(\rho)\mathcal{Q}_N(\delta)$, $\mathbf{P}_1^h(\delta) = \mathbf{C}_N'(\delta)\mathbb{L}_\lambda'(\delta)\mathcal{P}_N(\delta)$, $\mathbf{P}_2^h(\delta) = \mathbf{C}_N'(\delta)\mathbb{L}_\rho'(\rho)\mathcal{P}_N(\delta)$ and $\mathbf{P}_3^h(\delta) = \mathbf{C}_N'(\delta)\mathbb{N}_N'(\rho)\mathbb{L}_\rho'(\rho)\mathcal{Q}_N(\delta)$.

Note that $\mathbf{Q}_1^h(\delta) = \mathbf{C}_N'(\delta)[\bar{\mathbf{F}}_N'(\delta) - \bar{\mathbb{F}}_N'(\delta)]\mathcal{Q}_N(\delta) = \mathbf{W}'\mathbf{B}_N'(\rho)\mathcal{Q}_N(\delta) - \mathbf{C}_N'(\delta)\bar{\mathbb{F}}_N'(\delta)\mathcal{Q}_N(\delta)$. As $\bar{\mathbb{F}}_N'(\delta)$ is a diagonal matrix which is naturally bounded in both row and column sums, uniformly in $\delta \in \Delta$, we conclude $\mathbf{Q}_1^h(\delta)$ is bounded in both row and column sum norms, uniformly in $\delta \in \Delta$, by Lemma B.1. Similarly, $\mathbf{Q}_2^h(\delta) = \mathbf{C}_N'(\delta)\mathbb{M}_N'(\rho)[\bar{\mathbf{G}}_N(\rho) - \bar{\mathbb{G}}_N(\rho)]\mathcal{Q}_N(\delta) = \bar{\mathbf{Q}}(\delta) - \mathbf{C}_N'(\delta)\mathbb{M}_N'(\rho)\bar{\mathbb{G}}_N(\rho)\mathcal{Q}_N(\delta)$ is also bounded in both row and column sum norms, uniformly in $\delta \in \Delta$. Hence, $\mathbf{Q}_1^h(\delta)$ and $\mathbf{Q}_2^h(\delta)$ have forms similar to $\mathbf{Q}(\delta)$. The proof of (i) thus follows that of Theorem 1 **(b)**. For (ii), noting that $\mathcal{P}_N(\delta) = \mathbb{P}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{C}_N(\delta)$, we have $\sup_{\delta\in\Delta} \frac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}_s^h(\delta)\mathbf{C}_N^{-1}] = o(1), s = 1, 2$, by Lemma B.4. For the final result,

$$\frac{1}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}_3^h(\delta)\mathbf{C}_N^{-1}] = -\frac{1}{N_1}\mathtt{tr}[\mathbf{C}_N'(\delta)\mathbb{N}_N'(\rho)\mathbb{L}_\rho'(\rho)\mathcal{Q}_N(\delta)\mathrm{Var}(\mathbf{Y})]$$

$$= -\frac{1}{N_1}\mathtt{tr}\big[(\tfrac{1}{N_1}\mathbb{X}'(\rho)\mathbb{X}(\rho))^{-1}(\tfrac{1}{N_1}\mathbf{X}'\mathbf{B}_N'(\rho)\mathbb{L}_\rho'(\rho)\mathbb{Q}_{\mathbb{X}}(\rho)\mathbb{Q}_{\mathbb{D}}(\rho)\mathcal{C}_N^h(\delta)\mathbb{X}(\rho))\big].$$

Assumption C implies that the elements of $[\frac{1}{N_1}\mathbb{X}'(\rho)\mathbb{X}(\rho)]^{-1}$ are uniformly bounded for large enough N, uniformly in $\rho \in \Delta_\rho$. Lemma B.1 and Lemma B.3 together imply the term between $\mathbf{X}'$ and $\mathbb{X}(\rho)$ are uniformly bounded in both row and column sums, uniformly in $\delta \in \Delta$. Hence, the elements of the second part in the trace are also uniformly bounded. As the number of regressors k is finite, the quantity $\frac{\sigma_{v0}^2}{N_1}\mathtt{tr}[\mathbf{C}_N'^{-1}\mathbf{P}_3^h(\delta)\mathbf{C}_N^{-1}]$ will shrink to zero as N goes large, uniformly in $\delta \in \Delta$. These complete the proof. $\square$

Proof of Theorem 4. Applying the MVT on each row of $S_N^\diamond(\hat{\xi}_N^\diamond)$, we have,

$$0 = \frac{1}{\sqrt{N_1}}S_N^\diamond(\hat{\xi}_N^\diamond) = \frac{1}{\sqrt{N_1}}S_N^\diamond(\xi_0) + \big[\tfrac{1}{N_1}\tfrac{\partial}{\partial\xi'}S_N^\diamond(\xi)\big|_{\xi=\bar{\xi}_r \text{ in } r\text{th row}}\big]\sqrt{N_1}(\hat{\xi}_N^\diamond - \xi_0),$$

where $\{\bar{\xi}_r\}$ are on the line segment between $\hat{\xi}_N^\diamond$ and ξ_0. The result follows if

- (a) $\frac{1}{\sqrt{N_1}}S_N^\diamond(\xi_0) \xrightarrow{D} N[0,\ \lim_{N\to\infty}\Gamma_N^\diamond(\xi_0)]$,
- (b) $\frac{1}{N_1}\big[\frac{\partial}{\partial\xi'}S_N^\diamond(\xi)\big|_{\xi=\bar{\xi}_r \text{ in } r\text{th row}} - \frac{\partial}{\partial\xi'}S_N^\diamond(\xi_0)\big] = o_p(1)$, and
- (c) $\frac{1}{N_1}\big[\frac{\partial}{\partial\xi'}S_N^\diamond(\xi_0) - \mathrm{E}(\frac{\partial}{\partial\xi'}S_N^\diamond(\xi_0))\big] = o_p(1)$.

Proof of (a). It is easy to see that the elements of $S_N^\diamond(\xi_0)$ are linear-quadratic forms in $\mathbf{V}$. Thus, for every non-zero $(k+2)\times 1$ vector of constants a, $a'S_N^\diamond(\xi_0)$ has form:

$$a'S_N^\diamond(\xi_0) = b_N'\mathbf{V} + \mathbf{V}'\Phi_N\mathbf{V} - \sigma_v^2\mathtt{tr}(\Phi_N),$$

for suitably defined non-stochastic vector b_N and matrix Φ_N. By Assumptions A-F, we verify that b_N and matrix Φ_N satisfy the conditions of the CLT for LQ form of Kelejian and Prucha

(2001), and hence the asymptotic normality of $\frac{1}{\sqrt{N_1}} a' S^{\diamond}_N(\xi_0)$ follows. By Cramér-Wold device, $\frac{1}{\sqrt{N_1}} S^{\diamond}_N(\xi_0) \xrightarrow{D} N[0, \lim_{N\to\infty} \Gamma^{\diamond}_N(\theta_0)]$, where $\Gamma^{\diamond}_N(\theta_0)$ is in (4.5).

Proof of (b). The Hessian matrix $H^{\diamond}_N(\xi) = \frac{\partial}{\partial \xi'} S^{\diamond}_N(\xi)$ is given in Appendix A. As $\bar{\mathbb{F}}'_{N\lambda}(\delta_0)$, $\bar{\mathbb{F}}'_{N\rho}(\delta_0)$ and $\bar{\mathbb{G}}_{N\rho}(\rho_0)$ are diagonal matrices with uniformly bounded elements, we show that $\frac{1}{N_1} H^{\diamond}_N(\xi_0) = O_p(1)$ by Lemma B.6, and hence, $\frac{1}{N_1} H^{\diamond}_N(\bar{\xi}) = O_p(1)$. Here again we simply use $H^{\diamond}_N(\bar{\xi})$ to denote $\frac{\partial}{\partial \xi'} S^{\diamond}_N(\xi)\big|_{\xi=\bar{\xi}_r \text{ in } r\text{th row}}$. As $H^{\diamond}_N(\bar{\xi})$ is linear or quadratic in $\bar{\beta}$ and nonlinear in $\bar{\delta}$, we have by applying the MVT on the $\bar{\delta}$-components:

$$\tfrac{1}{N_1} H^{\diamond}_N(\bar{\xi}) - \tfrac{1}{N_1} H^{\diamond}_N(\xi_0) = \tfrac{1}{N_1} \tfrac{\partial}{\partial \delta'} H^{\diamond}_N(\bar{\beta}, \dot{\delta})(\bar{\delta} - \delta_0) + \tfrac{1}{N_1}[H^{\diamond}_N(\bar{\beta}, \delta_0) - H^{\diamond}_N(\theta_0)].$$

Similar to the proof of Theorem 2.2 (b), we show that $\frac{1}{N_1} \frac{\partial}{\partial \delta'} H^{\diamond}_N(\bar{\beta}, \dot{\delta}) = O_p(1)$. The second term is seen to contain elements either linear or quadratic in $\bar{\beta} - \beta_0$ with the matrices in the linear or quadratic terms being $O_p(1)$. Hence, the desired result follows as $\bar{\xi} - \xi_0 = o_p(1)$.

Proof of (c). Since $\mathbf{Y} = \mathbf{A}_N^{-1}(\eta + \mathbf{B}_N^{-1}\mathbf{V})$, all components of $H^{\diamond}_N(\xi_0)$ are linear or quadratic in $\mathbf{V}$. Thus, under the assumptions of the theorem, the result **(c)** is proved using Lemma B.6. We provide details of the proof using the most complicated term, $H^{\diamond}_{\rho\rho}(\xi_0)$. The proofs for the other terms are similar and thus are omitted.

Let $\Xi_N = -\mathbf{G}'_N \bar{\mathbb{G}}_N + \bar{\mathbb{G}}_{N\rho} + \bar{\mathbb{G}}_N \mathbb{G}_N$. By Lemma B.1, it is easy to see that Ξ_N is uniformly bounded in both row and column sums in absolute value. Hence, we have,

$$\begin{aligned}
&\tfrac{1}{N_1}[H^{\diamond}_{\rho_0\rho_0}(\xi_0) - \mathrm{E}(H^{\diamond}_{\rho_0\rho_0}(\xi_0))] \\
&= \tfrac{1}{N_1}[\mathbf{V}'\mathbb{Q}_{\mathbb{D}}\mathcal{R}_{1N}\mathbb{Q}_{\mathbb{D}}\mathbf{V} - \mathrm{E}(\mathbf{V}'\mathbb{Q}_{\mathbb{D}}\mathcal{R}_{1N}\mathbb{Q}_{\mathbb{D}}\mathbf{V})] - \tfrac{1}{N_1}(\mathbf{A}_N\mathbf{Y} - \mathbf{X}\beta_0)'\mathbf{B}'_N\Xi_N\mathbb{Q}_{\mathbb{D}}\mathbf{V} \\
&\quad + \tfrac{1}{N_1}\mathrm{E}[(\mathbf{A}_N\mathbf{Y} - \mathbf{X}\beta_0)'\mathbf{B}'_N\Xi_N\mathbb{Q}_{\mathbb{D}}\mathbf{V}] \\
&= \tfrac{1}{N_1}[\mathbf{V}'\mathbb{Q}_{\mathbb{D}}\mathcal{R}_{1N}\mathbb{Q}_{\mathbb{D}}\mathbf{V} - \mathrm{E}(\mathbf{V}'\mathbb{Q}_{\mathbb{D}}\mathcal{R}_{1N}\mathbb{Q}_{\mathbb{D}}\mathbf{V})] - \tfrac{1}{N_1}[\phi_0'\mathbb{D}'\Xi_N\mathbb{Q}_{\mathbb{D}}\mathbf{V} - \mathrm{E}(\phi_0'\mathbb{D}'\Xi_N\mathbb{Q}_{\mathbb{D}}\mathbf{V})] \\
&\quad - \tfrac{1}{N_1}[\mathbf{V}'\Xi_N\mathbb{Q}_{\mathbb{D}}\mathbf{V} - \mathrm{E}(\mathbf{V}'\Xi_N\mathbb{Q}_{\mathbb{D}}\mathbf{V})] \;=\; o_p(1),
\end{aligned}$$

completing the proof or Theorem 4. □

Corollary D.1. *Under the assumptions of Theorem 4, we have,*

$$\Gamma^{\diamond}_N(\hat{\xi}^{\diamond}_N, \hat{\phi}^{\diamond}_N, \mathbf{H}) = \Gamma^{\diamond}_N(\xi_0) + \text{Bias}^{\diamond}_{\phi}(\delta_0, \mathbf{H}) + o_p(1),$$

where $\text{Bias}^{\diamond}_{\phi}(\delta_0, \mathbf{H})$ *is a* $(k+2) \times (k+2)$ *matrix with all the* β*-related entries being zero and the* δ *entry of the elements:* $\frac{1}{N_1}\texttt{tr}(\mathbf{H}\mathbb{P}_{\mathbb{D}}\mathbb{L}'_a\mathbf{H}\mathbb{L}_b\mathbb{P}_{\mathbb{D}})$*, for* $a, b = \lambda, \rho$*.*

Proof: Just like the homoskedasiticiy case, plugging $\hat{\phi}^{\diamond}_N$ in $\Gamma^{\diamond}_N(\xi)$ induces a bias for terms quadratic in ϕ, and a bias correction is necessary. From (4.5), we see that only the (λ, ρ) components of $\Gamma^{\diamond}_N(\xi)$ contain terms quadratic in ϕ: $\phi'\mathbb{D}'(\rho)\mathbb{L}'_a(\delta)\mathbf{H}\mathbb{L}_b(\delta)\mathbb{D}(\rho)\phi$, $a, b = \lambda, \rho$.

Applying MVT on $\mathbf{D}\hat{\phi}^{\diamond}_N$ w.r.t. $\hat{\rho}^{\diamond}_N$, for $\dot{\rho}$ between $\hat{\rho}^{\diamond}_N$ and ρ_0, $\mathbf{D}\hat{\phi}^{\diamond}_N = \mathbf{D}\phi_0 + \mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{V} - \mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{B}_N[\mathbf{W}\mathbf{Y}(\hat{\lambda}^{\diamond}_N - \lambda_0) + \mathbf{X}(\hat{\beta}^{\diamond}_N - \beta_0)] - \mathbb{R}_N(\dot{\rho})[\mathbf{A}_N(\hat{\lambda}^{\diamond}_N)\mathbf{Y} - \mathbf{X}\hat{\beta}^{\diamond}_N](\hat{\rho}^{\diamond}_N - \rho_0)$. Thus,

$$\frac{1}{N_1}\hat{\phi}^{\diamond\prime}_N\mathbb{D}'(\hat{\rho}^{\diamond}_N)\mathbb{L}'_a(\hat{\delta}^{\diamond}_N)\mathbf{H}\mathbb{L}_b(\hat{\delta}^{\diamond}_N)\mathbb{D}(\hat{\rho}^{\diamond}_N)\hat{\phi}^{\diamond}_N = \frac{1}{N_1}\phi'_0\mathbb{D}'(\hat{\rho}^{\diamond}_N)\mathbb{L}'_a(\hat{\delta}^{\diamond}_N)\mathbf{H}\mathbb{L}_b(\hat{\delta}^{\diamond}_N)\mathbb{D}(\hat{\rho}^{\diamond}_N)\phi_0$$

$$+ \frac{1}{N_1}\mathbf{V}'\mathbb{P}_{\mathbb{D}}\mathbf{B}_N^{-1\prime}\mathbf{B}'_N(\hat{\rho}^{\diamond}_N)\mathbb{L}'_a(\hat{\delta}^{\diamond}_N)\mathbf{H}\mathbb{L}_b(\hat{\delta}^{\diamond}_N)\mathbf{B}_N(\hat{\rho}^{\diamond}_N)\mathbf{B}_N^{-1}\mathbb{P}_{\mathbb{D}}\mathbf{V} + o_p(1)$$

$$= \frac{1}{N_1}\phi'_0\mathbb{D}'\mathbb{L}'_a\mathbf{H}\mathbb{L}_b\mathbb{D}\phi_0 + \frac{1}{N_1}\texttt{tr}\big[\mathbf{H}\mathbb{P}_{\mathbb{D}}\mathbb{L}'_a\mathbf{H}\mathbb{L}_b\mathbb{P}_{\mathbb{D}}\big] + o_p(1),$$

after plugging $\mathbf{D}\hat{\phi}^{\diamond}_N$ and other parameter estimates. Therefore, $\frac{1}{N_1}\texttt{tr}\big[\mathbf{H}\mathbb{P}_{\mathbb{D}}(\rho)\mathbb{L}'_a(\delta)\mathbf{H}\mathbb{L}_b(\delta)\mathbb{P}_{\mathbb{D}}(\rho)\big]$, for $a, b = \lambda, \rho$, gives the non-zero elements in the bias matrix $\text{Bias}^{\diamond}_{\phi}(\delta_0, \mathbf{H})$ for $\Gamma^{\diamond}_N(\hat{\xi}^{\diamond}_N)$. □

Lemma D.1. *Assume* $\Pi_N(\rho) = [\mathbb{Q}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho)]^{-1}$ *exists for* ρ *in a neighborhood of* ρ_0*, and is bounded in both row and column sum norms. Let* $A_N = [a_{ij}]$ *and* $B_N = [b_{ij}]$ *be square matrices of dimension* N *with zero diagonals and bounded row and column sum norms. Let* $C_N = [c_{ij}]$ *be an* $N \times N$ *matrix with diagonal elements being uniformly bounded. We have,*

(*i*) $\frac{1}{N}\texttt{tr}(\widehat{\mathbf{H}}C_N) - \frac{1}{N}\texttt{tr}(\mathbf{H}C_N) = o_p(1)$,

(*ii*) $\frac{1}{N}\texttt{tr}(\widehat{\mathbf{H}}A_N\widehat{\mathbf{H}}B_N) - \frac{2}{N}\texttt{tr}((A_N \odot B_N)\Pi_N\Lambda(\mathbf{H})\Pi_N) - \frac{1}{N}\texttt{tr}(\mathbf{H}A_N\mathbf{H}B_N) = o_p(1)$,

where $\Pi_N = \Pi_N(\rho_0)$*,* $\Lambda(\mathbf{H}) = \{(q'_j\mathbf{H}q_k)^2\}_{j,k=1}^N$*, and* q'_j *is the* j*th row of* $\mathbb{Q}_{\mathbb{D}}$*.*

The assumptions on $\Pi_N(\rho)$ in Lemma D.1 always hold for a balanced panel, and typically hold for a general unbalanced panel. To see the invertibility, we have, $\mathbb{Q}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho) = I_N - 2I_N \odot \mathbb{P}_{\mathbb{D}}(\rho) + \mathbb{P}_{\mathbb{D}}(\rho) \odot \mathbb{P}_{\mathbb{D}}(\rho)$. By Schur's product theorem, the last term is positive semi-definite. In addition, when T is not too small, $I_N - 2I_N \odot \mathbb{P}_{\mathbb{D}}(\rho)$ is positive definite, because the diagonal elements of $\mathbb{P}_{\mathbb{D}}(\rho)$ are of order $O_p(1/T)$ (see the proof of Lemma B.3). These are mainly for theoretical considerations and in practice, a generalized inverse can simply be used. The bias term in Corollary D.1 needs a further correction when $\mathbf{H}$ is replaced by $\hat{\mathbf{H}}$ as it contains elements of the form $\mathtt{tr}(\mathbf{H}A_N\mathbf{H}B_N)$ with diagonal elements of A_N and B_N not strictly zero. However, the effect of non-zero diagonals is shown to be negligible due to the existence of a lower ranked matrix $\mathbb{P}_{\mathbb{D}}$ orthogonal to $\mathbb{Q}_{\mathbb{D}}$.

Proof of Lemma D.1: Using $\tilde{\mathbf{V}}(\xi) = \mathbb{Q}_{\mathbb{D}}(\rho)\mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y} - \mathbf{X}\beta]$ defined in (3.4), let $\tilde{\mathbf{V}} = \tilde{\mathbf{V}}(\xi_0)$ and $\hat{\mathbf{V}} = \tilde{\mathbf{V}}(\hat{\xi}_N^\diamond)$ and denote their elements by $\{\tilde{v}_j\}$ and $\{\hat{v}_j\}$, respectively. Following (C.7), we have $\hat{v}_j \equiv \tilde{v}_j(\hat{\xi}_N^\diamond) = \tilde{v}_j + \psi_j'(\hat{\xi}_N^\diamond - \xi_0) + o_p(\|\hat{\xi}_N^\diamond - \xi_0\|)$. In vector form,

$$\hat{\mathbf{V}} = \tilde{\mathbf{V}} + \Psi_N(\hat{\xi}_N^\diamond - \xi_0) + o_p(\|\hat{\xi}_N^\diamond - \xi_0\|),$$

where $\Psi_N = (\psi_1, \psi_2, \ldots, \psi_N)'$, with ψ_j being defined below (C.7).

Define $\dot{\Pi}_N(\rho) = \frac{\partial}{\partial\rho}\Pi_N(\rho) = -2\Pi_N(\rho)[\dot{\mathbb{Q}}_{\mathbb{D}}(\rho) \odot \mathbb{Q}_{\mathbb{D}}(\rho)]\Pi_N(\rho)$. It is easy to see that $\|\dot{\Pi}_N(\rho)\|_1$ and $\|\dot{\Pi}_N(\rho)\|_\infty$ are bounded in a neighborhood of ρ_0. Let Π_{jh} and $\dot{\Pi}_{jh}$ be the respective elements of Π_N and $\dot{\Pi}_N$. We have by the MVT, for each $j, h = 1, 2, \ldots, N$, $\Pi_{jh}(\hat{\rho}_N^\diamond) = \Pi_{jh} + \dot{\Pi}_{jh}(\bar{\rho})(\hat{\rho}_N^\diamond - \rho_0) = \Pi_{jh} + \dot{\Pi}_{jh}(\hat{\rho}_N^\diamond - \rho_0) + o_p(\|\hat{\rho}_N^\diamond - \rho_0\|)$, where $\bar{\rho}$ lies between $\hat{\rho}_N^\diamond$ and ρ_0. In matrix form,

$$\Pi_N(\hat{\rho}_N^\diamond) = \Pi_N + \dot{\Pi}_N(\hat{\rho}_N^\diamond - \rho_0) + o_p(\|\hat{\rho}_N^\diamond - \rho_0\|).$$

Define $\hat{h} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \ldots, \hat{\sigma}_N^2)' = \Pi_N(\hat{\rho}_N^\diamond)(\hat{\mathbf{V}} \odot \hat{\mathbf{V}})$ and $\tilde{h} = \Pi_N(\tilde{\mathbf{V}} \odot \tilde{\mathbf{V}})$. As the elements of $\tilde{\mathbf{V}}$

are $O_p(1)$, rows of Ψ_N are $O_p(1)$, elements of Π_N and $\dot{\Pi}_N$ are $O(1)$, and $\hat{\xi}_N^* - \xi_0 = O_p(\frac{1}{\sqrt{N_1}})$,

$$\hat{h} = \tilde{h} + 2\Pi_N(\tilde{\mathbf{V}} \odot \Psi_N(\hat{\xi}_N^\diamond - \xi_0)) + \dot{\Pi}_N(\tilde{\mathbf{V}} \odot \tilde{\mathbf{V}})(\hat{\rho}_N^\diamond - \rho_0) + o_p(\|\hat{\xi}_N^\diamond - \xi_0\|). \tag{D.2}$$

Proof of Result (i). Let $c_N = (c_{11}, \cdots, c_{NN})'$ and $h = (\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2)'$. We have,

$$\tfrac{1}{N}[\mathtt{tr}(\hat{\mathbf{H}}C_N) - \mathtt{tr}(\mathbf{H}C_N)] = \tfrac{1}{N}c_N'(\hat{h} - h) = \tfrac{1}{N}c_N'(\hat{h} - \tilde{h}) + \tfrac{1}{N}c_N'(\tilde{h} - h).$$

The result follows if both terms above are $o_p(1)$. For the first term, we have, using (D.2),

$$\tfrac{1}{N}\vec{c}_N'(\hat{h} - \tilde{h}) = \tfrac{2}{N}\vec{c}_N'\Pi_N(\tilde{\mathbf{V}} \odot \Psi_N(\hat{\xi}_N^\diamond - \xi_0)) + \tfrac{1}{N}\vec{c}_N'\dot{\Pi}_N(\tilde{\mathbf{V}} \odot \tilde{\mathbf{V}})(\hat{\rho}_N^\diamond - \rho_0) + o_p(\|\hat{\xi}_N^\diamond - \xi_0\|)$$
$$= \tfrac{2}{N}\textstyle\sum_{j=1}^N c_{jj}(\sum_{h=1}^N \Pi_{jh}\tilde{v}_h\psi_h')(\hat{\xi}_N^\diamond - \xi_0) + \tfrac{1}{N}\sum_{j=1}^N c_{jj}(\sum_{h=1}^N \dot{\Pi}_{jh}\sum_{k=1}^N q_{hk}^2\sigma_k^2)(\hat{\rho}_N^\diamond - \rho_0)$$
$$+\ o_p(\|\hat{\xi}_N^\diamond - \xi_0\|) = o_p(1).$$

For the second term, we have after some algebra,

$$\tilde{h} = \Pi_N[(\mathbb{Q}_{\mathbb{D}} \odot \mathbb{Q}_{\mathbb{D}})(\mathbf{V} \odot \mathbf{V}) + \zeta] = \mathbf{V} \odot \mathbf{V} + \Pi_N\varepsilon, \tag{D.3}$$

where ε is an $N\times 1$ vector with j-th element $\varepsilon_j = \sum_{k=1}^N v_k\zeta_{jk}$, where $\zeta_{jk} = 2q_{jk}\sum_{l=1}^{k-1} q_{jl}v_l, k \geq 2$, and $\zeta_{j1} = 0$. As ζ_{jk} is $(v_1, \dots v_{k-1})$-measurable, $\{v_k\zeta_{jk}\}$ form an M.D. sequence. Thus, each ε_j is a sum of M.D.s. Hence, we have

$$\tfrac{1}{N}\vec{c}_N'(\tilde{h} - h) = \tfrac{1}{N}\vec{c}_N'(\mathbf{V} \odot \mathbf{V} - h) + \tfrac{1}{N}\vec{c}_N'\Pi_N\zeta = o_p(1),$$

by Lemma B.6(v) and WLLN of Davidson (1994, Theorem 19.7) for M.D. arrays.

Proof of Result (ii). Note that $\mathtt{tr}(\mathbf{H}A_N\mathbf{H}B_N) = h'(A_N \odot B_N)h$. We have,

$$\tfrac{1}{N}\mathtt{tr}(\widehat{\mathbf{H}}A_N\widehat{\mathbf{H}}B_N) - \tfrac{1}{N}\mathtt{tr}(\mathbf{H}A_N\mathbf{H}B_N) = \tfrac{1}{N}\hat{h}'(A_N \odot B_N)\hat{h} - \tfrac{1}{N}h'(A_N \odot B_N)h$$
$$= \tfrac{1}{N}(\hat{h}'(A_N \odot B_N)\hat{h} - \tilde{h}'(A_N \odot B_N)\tilde{h}) + \tfrac{1}{N}(\tilde{h}'(A_N \odot B_N)\tilde{h} - h'(A_N \odot B_N)h). \tag{D.4}$$

The first term of (D.4) can be written as

$$\tfrac{1}{N}(\hat{h}'(A_N \odot B_N)\hat{h} - \tilde{h}'(A_N \odot B_N)\tilde{h}) = \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3,$$

where $\mathcal{T}_1 = \frac{1}{N}(\hat{h} - \tilde{h})'(A_N \odot B_N)(\hat{h} - \tilde{h})$, $\mathcal{T}_2 = \frac{1}{N}(\hat{h} - \tilde{h})'(A_N \odot B_N)\tilde{h}$, and $\mathcal{T}_3 = \frac{1}{N}(\hat{h} -$

$\tilde{h})'(A_N \odot B_N)'\tilde{h}$. As A_N and B_N are uniformly bounded in both row and column sum norms, $A_N \odot B_N$ is also uniformly bounded in both row and column sum norms. Hence, using (D.2), $\tilde{\mathbf{V}} = O_p(1)$, $\Psi_N = O_p(1)$ and $\hat{\xi}_N^* - \xi_0 = O_p(\frac{1}{\sqrt{N_1}})$, we show that $\mathcal{T}_r = o_p(1), r = 1, 2, 3$, in a way similar to the proof of $\frac{1}{N}\vec{c}'_N(\hat{h} - \tilde{h}) = o_p(1)$ in the proof of (i) above. Thus, the first term in (D.4) is $o_p(1)$.

For the second term in (D.4), we have similarly to the first term,

$$\tfrac{1}{N}(\tilde{h}'(A_N \odot B_N)\tilde{h} - h'(A_N \odot B_N)h) = \mathcal{T}_4 + \mathcal{T}_5 + \mathcal{T}_6,$$

where $\mathcal{T}_4 = \frac{1}{N}(\tilde{h}-h)'(A_N \odot B_N)(\tilde{h}-h)$, $\mathcal{T}_5 = \frac{1}{N}(\tilde{h}-h)'(A_N \odot B_N)h$, and $\mathcal{T}_6 = \frac{1}{N}(\tilde{h}-h)'(A_N \odot B_N)'h$. For $\mathcal{T}_5$ term, we have by (D.3), $\mathcal{T}_5 = \frac{1}{N}(\mathbf{V}\odot\mathbf{V}-h)'(A_N \odot B_N)h + \frac{1}{N}\varepsilon'\Pi_N(A_N \odot B_N)h = o_p(1)$, by Lemma B.6$(v)$ and WLLN for M.D. arrays of Davidson (1994, Theorem 19.7). The $\mathcal{T}_6$ term is similar to $\mathcal{T}_5$ and the result follows, i.e., $\mathcal{T}_6 = o_p(1)$. Finally for $\mathcal{T}_4$, we have again,

$$\mathcal{T}_4 = \mathcal{T}_{4a} + \mathcal{T}_{4b} + \mathcal{T}_{4c} + \mathcal{T}_{4d} \tag{D.5}$$

by (D.3), where $\mathcal{T}_{4a} = \frac{1}{N}(\mathbf{V} \odot \mathbf{V} - h)'(A_N \odot B_N)\Pi_N\varepsilon$, $\mathcal{T}_{4b} = \frac{1}{N}(\mathbf{V} \odot \mathbf{V} - h)'(A_N \odot B_N)'\Pi_N$, $\mathcal{T}_{4c} = \frac{1}{N}(\mathbf{V} \odot \mathbf{V} - h)'(A_N \odot B_N)(\mathbf{V} \odot \mathbf{V} - h)$, and $\mathcal{T}_{4d} = \frac{1}{N}\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N$. Consider first the term $\mathcal{T}_{4a}$. Denote $\Omega = (A_N \odot B_N)\Pi_N$ with elements $\{\omega_{jk}\}$. We have,

$$\mathcal{T}_{4a} = \tfrac{1}{N}\textstyle\sum_{j=1}^N \sum_{k=1}^N \omega_{jk}\varepsilon_j(v_k^2 - \sigma_k^2) = \tfrac{1}{N}\sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N \sum_{\substack{m\neq l\\ m=1}}^N \omega_{jk}q_{jl}q_{jm}(v_k^2 - \sigma_k^2)v_l v_m,$$

which is further decomposed into:

$$\begin{aligned}
&\tfrac{1}{N}\textstyle\sum_{k=1}^N((v_k^2 - \sigma_k^2)\sum_{j=1}^N \sum_{l=1}^{k-1} \sum_{\substack{m\neq l\\ m=1}}^{k-1} \omega_{jk}q_{jl}q_{jm}v_l v_m)\\
&+ \tfrac{2}{N}\textstyle\sum_{l=1}^N(v_l \sum_{j=1}^N \sum_{k=1}^{l-1} \sum_{\substack{m\neq k\\ m=1}}^{l-1} \omega_{jk}q_{jl}q_{jm}v_m(v_k^2 - \sigma_k^2))\\
&+ \tfrac{2}{N}\textstyle\sum_{k=1}^N((v_k^3 - \mathrm{E}v_k^3)\sum_{j=1}^N \sum_{m=1}^{k-1} \omega_{jk}q_{jk}q_{jm}v_m)\\
&+ \tfrac{2}{N}\textstyle\sum_{m=1}^N(v_m \sum_{j=1}^N \sum_{k=1}^{m-1} \omega_{jk}q_{jk}q_{jm}(v_k^3 - \mathrm{E}v_k^3))\\
&+ \tfrac{2}{N}\textstyle\sum_{m=1}^N(v_m \sum_{j=1}^N \sum_{\substack{k\neq m\\ k=1}}^N \omega_{jk}q_{jk}q_{jm}(\mathrm{E}v_k^3 - \sigma_k^2)).
\end{aligned}$$

Each term is the average of an M.D. sequence and thus is $o_p(1)$ by Theorem 19.7 of Davidson (1994). Similarly, we show that $\mathcal{T}_{4b} = \frac{1}{N}(\mathbf{V} \odot \mathbf{V} - h)'(A_N \odot B_N)'\Pi_N \varepsilon = o_p(1)$.

For the term $\mathcal{T}_{4c}$, as $\mathrm{E}(\mathbf{V} \odot \mathbf{V}) = h$, we have $\mathrm{E}(\mathcal{T}_{4c}) = \frac{1}{N}\mathtt{tr}((A_N \odot B_N)\mathrm{Var}(\mathbf{V} \odot \mathbf{V})) = 0$. Thus, Lemma B.6$(iv)$ implies that $\mathcal{T}_{4c} = \frac{1}{N}(\mathbf{V} \odot \mathbf{V} - h)'(A_N \odot B_N)(\mathbf{V} \odot \mathbf{V} - h) \xrightarrow{p} 0$.

Now, for the last term of (D.5), $\mathcal{T}_{4d} = \frac{1}{N}\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon$, we have by taking the advantage that each element of ε is a sum of an M.D. sequence,

$$\mathrm{E}(\varepsilon\varepsilon') = 2(\mathbb{Q}_{\mathbb{D}}\mathbf{H}\mathbb{Q}_{\mathbb{D}}) \odot (\mathbb{Q}_{\mathbb{D}}\mathbf{H}\mathbb{Q}_{\mathbb{D}}) - 2(\mathbb{Q}_{\mathbb{D}} \odot \mathbb{Q}_{\mathbb{D}})\mathbf{H}\mathbf{H}(\mathbb{Q}_{\mathbb{D}} \odot \mathbb{Q}_{\mathbb{D}}). \tag{D.6}$$

Let $\Lambda(\mathbf{H}) = (\mathbb{Q}_{\mathbb{D}}\mathbf{H}\mathbb{Q}_{\mathbb{D}}) \odot (\mathbb{Q}_{\mathbb{D}}\mathbf{H}\mathbb{Q}_{\mathbb{D}})$. As A_N and B_N have zero diagonal elements, we have

$$\begin{aligned} \mathrm{E}\big[\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon\big] &= 2\mathtt{tr}\big[(A_N \odot B_N)\Pi_N\Lambda(\mathbf{H})\Pi_N\big] - 2\mathtt{tr}\big[(A_N \odot B_N)\mathbf{H}^2\big], \qquad \text{(D.7)} \\ &= 2\mathtt{tr}\big[(A_N \odot B_N)\Pi_N\Lambda(\mathbf{H})\Pi_N\big]. \end{aligned}$$

Finally, to show that $\mathcal{T}_{4d} - \mathrm{E}(\mathcal{T}_{4d}) = o_p(1)$, denote $\chi_N = \Pi_N(A_N \odot B_N)\Pi_N$ with elements $\{\chi_{jk}\}$. It is easy to show that $\{\chi_{jk}\}$ are uniformly bounded. Let $|\chi_{lm}| \le \bar{\chi} < \infty$. We have,

$$\begin{aligned} &\mathrm{Var}(\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon) \\ &= 8\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N\sum_{m=1}^N\sum_{h=1}^N\sum_{\substack{p\neq h\\p=1}}^N\sum_{s=1}^N\sum_{\substack{r\neq s\\r=1}}^N \chi_{jk}\chi_{lm}q_{jh}q_{jp}q_{lh}q_{lp}q_{ks}q_{kr}q_{ms}q_{mr}\mathrm{E}(v_h^2v_p^2v_s^2v_r^2) \\ &\le 8\bar{q}^2\bar{\chi}c\sum_{m=1}^N(\sum_{j=1}^N|\chi_{jk}|)(\sum_{k=1}^N|q_{kr}|)(\sum_{l=1}^N|q_{lp}|)(\sum_{h=1}^N|q_{lh}|)(\sum_{p=1}^N|q_{jp}|)(\sum_{s=1}^N|q_{ms}|)(\sum_{r=1}^N|q_{mr}|) \end{aligned}$$

which is seen to be $= O(N)$. The inequality holds because $\mathrm{E}(v_h^2v_p^2v_s^2v_r^2)$ equals either $\mathrm{E}(v_h^2v_s^2)\mathrm{E}(v_p^2v_r^2)$ or $\mathrm{E}(v_h^2v_r^2)\mathrm{E}(v_p^2v_s^2)$ since $h \neq p$ and $s \neq r$. Both terms are bounded since, e.g., $\mathrm{E}(v_h^2v_r^2) \le \mathrm{E}^{\frac{1}{2}}(v_h^4)\mathrm{E}^{\frac{1}{2}}(v_r^4) \le c < \infty$. Therefore, by Chebyshev's inequality,

$$\begin{aligned} &P(\tfrac{1}{N}|\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon - \mathrm{E}(\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon)| \ge M) \\ &\le \tfrac{1}{M^2}\tfrac{1}{N^2}\mathrm{Var}(\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon) = o(1). \end{aligned}$$

It follows that $\frac{1}{N}\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon - \frac{1}{N}\mathrm{E}(\varepsilon'\Pi_N(A_N \odot B_N)\Pi_N\varepsilon) \xrightarrow{p} 0$. Therefore, we have shown that $\mathcal{T}_4 = \frac{2}{N}\mathtt{tr}((A_N \odot B_N)\Pi_N\Lambda(\mathbf{H})\Pi_N) + o_p(1)$. It follows that

$$\tfrac{1}{N}\mathtt{tr}(\widehat{\mathbf{H}}A_N\widehat{\mathbf{H}}B_N^{\circ}) - \tfrac{1}{N}\mathtt{tr}(\mathbf{H}A_N\mathbf{H}B_N^{\circ}) = \textstyle\sum_{r=1}^{6}\mathcal{T}_r = \tfrac{2}{N}\mathtt{tr}((A_N \odot B_N)\Pi_N\Lambda(\mathbf{H})\Pi_N) + o_p(1),$$

completing the proof of Lemma D.1. □

Corollary D.2. *Under the assumptions of Theorem 4, as $N \to \infty$, $\widehat{\Sigma}_N^{\diamond} - \Sigma_N^{\diamond}(\xi_0) \xrightarrow{p} 0$ and $\widehat{\Gamma}_N^{\diamond} - \Gamma_N^{\diamond}(\xi_0) \xrightarrow{p} 0$. Therefore, $\widehat{\Sigma}_N^{\diamond-1}\widehat{\Gamma}_N^{\diamond}\widehat{\Sigma}_N^{\diamond\prime-1} - \Sigma_N^{\diamond-1}(\xi_0)\Gamma_N^{\diamond}(\xi_0)\Sigma_N^{\diamond\prime-1}(\xi_0) \xrightarrow{p} 0$.*

Proof: The consistency of $\widehat{\Sigma}_N^{\diamond}$ to $\Sigma_N^{\diamond}(\xi_0)$ is implied by results (b) and (c) in the proof of Theorem 4. To show $\widehat{\Gamma}_N^{\diamond} - \Gamma_N^{\diamond}(\xi_0) \xrightarrow{p} 0$, it is easy to argue that:

- transition from $\Gamma_N^{\diamond}(\xi_0, \phi_0, \mathbf{H})$ to $\Gamma_N^{\diamond}(\hat{\xi}_N^{\diamond}, \phi_0, \mathbf{H})$ is asymptotically costless;
- cost of transition from $\Gamma_N^{\diamond}(\hat{\xi}_N^{\diamond}, \phi_0, \mathbf{H})$ to $\Gamma_N^{\diamond}(\hat{\xi}_N^{\diamond}, \hat{\phi}_N^{\diamond}, \mathbf{H})$ is captured by $\text{Bias}_{\phi}^{\diamond}(\hat{\delta}_N^{\diamond}, \mathbf{H})$;
- effect of replacing $\mathbf{H}$ by $\hat{\mathbf{H}}$ in $\frac{1}{N_1}\mathtt{tr}(\mathbf{H}\mathbb{L}_a\mathbf{H}\mathbb{L}_b^{\circ})$ is captured by $\frac{2}{N_1}\mathtt{tr}((\mathbb{L}_a \odot \mathbb{L}_b^{\circ})\Pi_N\Lambda(\mathbf{H})\Pi_N)$;

for $a, b = \lambda, \rho$. It is left to show that the cost of transition from $\text{Bias}_{\phi}^{\diamond}(\hat{\delta}_N^{\diamond}, \mathbf{H})$ to $\text{Bias}_{\phi}^{\diamond}(\hat{\delta}_N^{\diamond}, \hat{\mathbf{H}})$ is captured by $-\frac{2}{N_1}\mathtt{tr}((\mathbb{P}_{\mathbb{D}}\mathbb{L}_a' \odot \mathbb{L}_b\mathbb{P}_{\mathbb{D}})\Pi_N\Lambda(\mathbf{H})\Pi_N), a, b = \lambda, \rho$.

The non-zero entries in $\text{Bias}_{\phi}^{\diamond}(\delta_0, \mathbf{H})$ are of the form $\frac{1}{N_1}\mathtt{tr}(\mathbf{H}\mathbb{P}_{\mathbb{D}}\mathbb{L}_a'\mathbf{H}\mathbb{L}_b\mathbb{P}_{\mathbb{D}})$, for $a, b = \lambda, \rho$, as given in Corollary D.1. Applying result (D.7) with $A_N = \mathbb{P}_{\mathbb{D}}\mathbb{L}_a'$ and $B_N = \mathbb{L}_b\mathbb{P}_{\mathbb{D}}$, we have,

$$\begin{aligned}
&\tfrac{1}{N_1}\mathtt{tr}\big[\mathbb{P}_{\mathbb{D}}(\hat{\rho}_N^{\diamond})\mathbb{L}_a'(\hat{\delta}_N^{\diamond})\hat{\mathbf{H}}\mathbb{L}_b(\hat{\delta}_N^{\diamond})\mathbb{P}_{\mathbb{D}}(\hat{\rho}_N^{\diamond})\hat{\mathbf{H}} - \mathbb{P}_{\mathbb{D}}\mathbb{L}_a'\mathbf{H}\mathbb{L}_b\mathbb{P}_{\mathbb{D}}\mathbf{H}\big] \\
=&\tfrac{1}{N_1}\mathtt{tr}\big[\mathbb{P}_{\mathbb{D}}\mathbb{L}_a'\hat{\mathbf{H}}\mathbb{L}_b\mathbb{P}_{\mathbb{D}}\hat{\mathbf{H}} - \mathbb{P}_{\mathbb{D}}\mathbb{L}_a'\mathbf{H}\mathbb{L}_b\mathbb{P}_{\mathbb{D}}\mathbf{H}\big] + o_p(1) \ \ \text{(by the MVT)} \\
=&\tfrac{2}{N_1}\mathtt{tr}\big[(\mathbb{P}_{\mathbb{D}}\mathbb{L}_a' \odot \mathbb{L}_b\mathbb{P}_{\mathbb{D}})\Pi_N\Lambda(\mathbf{H})\Pi_N\big] + \tfrac{1}{N_1}\mathtt{tr}\big[(\mathbb{P}_{\mathbb{D}}\mathbb{L}_a' \odot \mathbb{L}_b\mathbb{P}_{\mathbb{D}})\mathbf{H}^2\big] + o_p(1),
\end{aligned}$$

for $a, b = \lambda, \rho$. Although the diagonal elements of $\mathbb{P}_{\mathbb{D}}\mathbb{L}_a' \odot \mathbb{L}_b\mathbb{P}_{\mathbb{D}}$ are not zero identically, they are shown to be small under Assumption F and therefore $\frac{1}{N_1}\mathtt{tr}((\mathbb{P}_{\mathbb{D}}\mathbb{L}_a' \odot \mathbb{L}_b\mathbb{P}_{\mathbb{D}})\mathbf{H}^2)$ is $o(1)$. The detail is tedious and thus is omitted. The result of Corollary D.2 thus follows. □

Appendix E: The Full Set of Monte Carlo Results

These Monte Carlo experiments are conducted to investigate (i) the finite sample performance of the proposed M-estimators and the corresponding standard error estimators, and (ii) the consequence of choosing a wrong estimator, i.e., choosing naïve or QML estimator instead of the proposed M-estimator. The `naïve` estimator is the M-estimator based on the balanced panel obtained by deleting units without full presence.

E.1 First-order Models

We consider two data generating processes (DGP): `GU-SPD-FE` models with `SL` (spatial lag) and `SE` (spatial error) or with `SL` and `SD` (spatial Durbin). The DGP with all the spatial effects (`SL`, `SE` and `SD`) is not considered due to the potential identification issues (Anselin et al., 2008 and Lee and Yu, 2016). For $t = 1, \ldots, T$,

$$\texttt{DGP 1}: \; Y_t = \lambda W_t Y_t + X_t \beta_1 + D_t \mu + \alpha_t l_{n_t} + U_t, \quad U_t = \rho M_t U_t + V_t,$$

$$\texttt{DGP 2}: \; Y_t = \lambda W_t Y_t + X_t \beta_1 + W_t X_t \beta_2 + D_t \mu + \alpha_t l_{n_t} + V_t.$$

To be self-contained, the descriptions given in the main text are repeated and more added. We choose $n = 50, 100, 200, 400$, and $T = 5, 10$. The parameters values are set at $\lambda = 0.2$, $\beta_1 = 1$, $\rho = 0.2$ for `DGP 1` and $\lambda = 0.2$, $\beta_1 = 1$, $\beta_2 = 0.5$ for `DGP 2`. $X_t's$ are generated independently from $N(0, 2^2 I_n)$, and individual effects are set to be $\mu = \frac{1}{T}\Sigma_{t=1}^T X_t + e$, where $e \sim N(0, I_n)$. The time fixed effects α are generated from $N(0, I_T)$. The number of Monte Carlo runs is 1000.

The spatial weight matrices can be `Rook` contiguity, `Queen` contiguity, or `Group` interaction. To generate W_t under `Rook`, randomly permute the indices $\{1, 2, \ldots, n\}$ for n spatial units and then allocate them into a lattice of $k \times m$ squares. Let $W_{nt,ij} = 1$ if the index j is in a square that is immediately left or right, above, or below the square that contains the index i. Similarly, W_{nt} under `Queen` is generated with additional neighbors sharing a common vertex with the unit i. To generate W_{nt} under `Group`, let $W_{nt,ij} = 1$ if units i and j belong to the same group. The distribution of the idiosyncratic errors $\{v_{it}\}$ can be (i) `normal`, (ii) standardized `normal mixture` (10% $N(0, 4^2)$ and 90% $N(0, 1)$), or (iii) standardized `chi-square` with 3 degrees of freedom. See Yang (2015) for details.

We consider both homoskedasticity and heteroskedasticity cases, with $\sigma_v^2 = 1$ for the homoskedastic case and the average of error variances 1 for the heteroskedastic. Two types of `Group` interaction weights are used: the groups' sizes are increasing with n with the number of groups $K(n) = Round(n^{0.5})$ (`Group-I`), or fixed by starting with six groups of fixed sizes (3, 5, 7, 9, 11, 15) and then replicating (`Group-II`). In the latter case, the variation in group sizes does not shrink to zero as n increases. As a result, the M-estimation would not be consistent under heteroskedasticity (Liu and Yang, 2015, 2020). In this case, the heteroskedasticity is generated as follows: for each group, if the group size is larger than the mean group size, then the variance is set to be the same as the group size, otherwise, the variance is the square of the inverse of the group size (Lin and Lee, 2010).

The selection matrices D_t in `DGP 1` and `DGP 2` are generated as follows: for each t, associate with each row of I_n a uniform (0, 1) random number, and the rows with random numbers smaller than $p_t \in (0, 1)$ are deleted, which corresponds to the $100p_t$% non-presence units. We fix the overall unbalancedness percentage at 10%. To generate spatial panel data with GU, we first generate the full vectors/matrices $(V_t^*, \mu, X_t^*, W_t^*, M_t^*)$ for each t, then do deletions according to the generated D_t to give $V_t = D_t V_t^*$, $X_t = D_t X_t^*$, $W_t = D_t W_t^* D_t'$, and $M_t = D_t M_t^* D_t'$, and then generate Y_t according to `DGP 1` or `DGP 2`. Monte Carlo (empirical) means and standard deviations (sd, shown in parentheses) are recorded for the `naïve` estimator, `QMLE`, M-estimator (`M-Est`), and RM-estimator (`RM-Est`). The empirical averages of the standard error estimates ($\hat{se}$, shown in square brackets) are also recorded for the `naïve` estimator, `M-Est` and `RM-Est`, based on the VC matrix estimates introduced in Sections 3 and 4.

Tables 1a and 1b report Monte Carlo results for the `GU-USPD-FE` model with `SL` and `SE` effects and homoscedastic errors, for different spatial weights combinations. The results show an excellent finite performance of the proposed `M-Est` and `RM-Est`, and their standard error estimators. The proposed `M-Est` performs uniformly much better than the `QMLE` in the estimation of σ_v^2, λ, and ρ, irrespective of the values of n and T. Our `M-Est` exhibits a good performance even when the sample size is as small as $n = 50$ and $T = 5$, and improves on average when the sample expands, regardless of the error distributions. Nonnormality does not have much effect on the performance of the estimators. The $\sqrt{N_1}$-consistency of the `M-Est` is clearly demonstrated by the Monte Carlo sds. Moreover, the robust estimates of standard errors $\hat{sd}$'s are on average very close to the corresponding Monte Carlo standard errors. By comparing the results of `M-Est` and `RM-Est`, we cannot see which one beats the other in terms of bias and efficiency for these homoscedastic models. By comparing the results of Table 1a and Table 1b, we see that a denser spatial weight matrix (e.g., `W = Group-I` vs `W = Rook`) can slow down the convergence.

Table 1c reports Monte Carlo results for the `naïve` estimator for the model with `SL` and `SE` effects and homoskedastic errors. The results show that the `naïve` estimators for spatial parameters are uniformly more biased than the other estimators, regardless of the sample size, as the connectivity structures of those deleted units are totally ignored.

Tables 2a and 2b present Monte Carlo results for the `GU-USPD-FE` model with `SL` and `SD` effects and homoskedastic errors, for W being `Queen` and `Group-I`, respectively. The results again show an excellent performance of the proposed set of estimation and inference methods. As in the case of the `SL-SE` model, the `M-Est` and `RM-Est` give quite similar results, and both show a clear convergence as the sample size increases. Their corresponding standard error estimates also perform very well. In contrast, the `QMLE` can perform poorly.

Tables 3a and 3b report Monte Carlo results for the model with `SL-SE` effects and heteroskedastic errors, for different spatial weights. The results show an excellent finite sample performance of the proposed `RM-Est` and their estimated standard errors. In contrast, the `QMLE` and `M-Est` typically provide worse estimates for spatial parameters than `RM-Est`. Our `RM-Est`s perform well even when the sample size is quite small, and show convergence to the true value as the sample size increases. In addition, $\hat{sd}$s are very close to sds for our `RM-Est`, consistent with our theoretical expectation. Together with simulation results in Tables 1a and 1b, we conclude that in real applications, when homoskedasticity holds either `M-Est` or `RM-Est` can be used, but when it is in doubt one should use the `RM-Est`.

Table 4 presents Monte Carlo results for the model with `SL-SD` effects and heteroskedastic errors. The weight matrix is specified as `Group-II`. The results show a much better finite sample performance of the `RM-Est` than `QMLE` and `M-Est` in terms of bias and standard error.

E.2 Second-order Models

Following are Monte Carlo results for high-order models ($p = q = 2$). The data-generating processes (DGP) we consider is

$$\texttt{DGP 3}: \; Y_t = \lambda_1 W_{1t} Y_t + \lambda_2 W_{2t} Y_t + X_t \beta_1 + D_t \mu + \alpha_t l_{n_t} + U_t, \quad U_t = \rho_1 M_{1t} + \rho_2 M_{2t} U_t + V_t,$$

for $t = 1, \ldots, T$. We choose $\lambda_1 = \rho_1 = 0.2$, $\lambda_2 = \rho_2 = 0.3$. The other parameters and the combination of n and T are set the same as `DGP 1` in section .

For homoskedasticity model, we set $\{W_{1t}\}$ and $\{M_{1t}\}$ as time-varying `Queen` contiguity, and $\{W_{2t}\}$ and $\{M_{2t}\}$ time-varying `Rook` contiguity. For heteroskedasticity model, we generate $\{W_{1t}\}$ and $\{M_{1t}\}$ based on time-varying `Group-II` interaction, and $\{W_{2t}\}$ and $\{M_{2t}\}$ time-varying `Rook` contiguity. All the other settings and the way of generating spatial panel data with GU are the same as first order models.

Table 5 reports Monte Carlo results for the second-order `GU-USPD-FE` model with `SL` and `SE` effects and homoscedastic errors. The results based on `M-Est` and `RM-Est` are still very close, and both of them perform uniformly much better than the `QMLE` in the estimation of spatial parameters and variance parameters for all the combinations of n and T. This holds for all of the error distributions. In addition, the robust estimates of standard errors $\hat{sd}$'s are uniformly close to the corresponding empirical standard errors.

Table 6 reports Monte Carlo results for the second-order model with `SL-SE` effects and heteroskedastic errors. The results also show an excellent finite sample performance of the proposed `RM-Est` and their estimated standard errors. The convergence to the true value as the sample size increases is clear for the proposed `RM-Est`. In contrast, the `QMLE` and `M-Est` can perform poorly even when the sample size is large enough. In addition, $\hat{sd}$s are very close to sds for our `RM-Est`, which is in line with our theoretical expectation. Again, one should use the `RM-Est` in practice whenever homoskedasticity is in doubt.

Table 1a: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP1` with **homoskedasticity**, Unbalancedness percentage = 10%, $(\beta, \lambda, \rho, \sigma_v^2) = (1, 0.2, 0.2, 1)$, and `W = Rook` and `M = Queen`.

	T=5			T=10		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
			$n = 50$; error = 1, 2, 3, for the three panels below			
β_1	.9998(.039)	1.0007(.039)[.039]	1.0007(.039)[.039]	1.0019(.026)	1.0008(.026)[.026]	1.0008(.026)[.026]
λ	.1848(.063)	.1999(.063)[.062]	.1999(.063)[.062]	.1820(.039)	.1976(.039)[.040]	.1976(.039)[.040]
ρ	.1112(.152)	.1868(.146)[.148]	.1867(.146)[.146]	.1239(.093)	.1974(.091)[.091]	.1974(.092)[.091]
σ_v^2	.7394(.083)	.9829(.110)[.107]	–	.8641(.062)	.9930(.071)[.071]	–
β_1	.9981(.038)	.9989(.038)[.039]	.9989(.038)[.038]	1.0006(.026)	.9995(.026)[.026]	.9995(.026)[.026]
λ	.1849(.061)	.1998(.061)[.062]	.1999(.061)[.060]	.1849(.039)	.2004(.039)[.040]	.2004(.039)[.039]
ρ	.1179(.149)	.1933(.143)[.148]	.1932(.144)[.140]	.1203(.093)	.1941(.091)[.092]	.1940(.091)[.089]
σ_v^2	.7358(.172)	.9780(.228)[.215]	–	.8625(.144)	.9912(.166)[.156]	–
β_1	.9981(.038)	.9990(.038)[.039]	.9990(.038)[.038]	1.0019(.026)	1.0008(.026)[.026]	1.0008(.026)[.026]
λ	.1825(.061)	.1976(.061)[.062]	.1976(.061)[.061]	.1819(.040)	.1976(.040)[.040]	.1976(.040)[.040]
ρ	.1165(.150)	.1919(.144)[.148]	.1917(.144)[.143]	.1194(.094)	.1931(.093)[.091]	.1931(.093)[.090]
σ_v^2	.7421(.128)	.9864(.169)[.161]	–	.8667(.105)	.9962(.121)[.113]	–
			$n = 100$; error = 1, 2, 3, for the three panels below			
β_1	1.0010(.027)	1.0011(.026)[.027]	1.0011(.026)[.027]	1.0001(.018)	.9997(.018)[.018]	.9997(.018)[.018]
λ	.1922(.043)	.1993(.043)[.042]	.1994(.043)[.042]	.1924(.027)	.1993(.027)[.027]	.1993(.027)[.027]
ρ	.1565(.099)	.1906(.096)[.100]	.1906(.096)[.099]	.1600(.063)	.1952(.062)[.063]	.1952(.062)[.063]
σ_v^2	.7617(.060)	.9942(.078)[.076]	–	.8792(.044)	.9986(.050)[.050]	–
β_1	.9993(.028)	.9994(.028)[.027]	.9994(.028)[.027]	1.0005(.018)	1.0000(.018)[.018]	1.0000(.018)[.018]
λ	.1923(.042)	.1994(.042)[.042]	.1994(.042)[.042]	.1932(.027)	.2001(.027)[.027]	.2000(.027)[.027]
ρ	.1623(.102)	.1962(.099)[.099]	.1962(.099)[.096]	.1634(.062)	.1985(.061)[.063]	.1985(.061)[.062]
σ_v^2	.7624(.128)	.9951(.167)[.160]	–	.8773(.102)	.9964(.116)[.112]	–
β_1	.9983(.027)	.9984(.027)[.027]	.9984(.027)[.027]	1.0005(.018)	1.0001(.018)[.018]	1.0001(.018)[.018]
λ	.1937(.043)	.2009(.043)[.042]	.2009(.043)[.042]	.1923(.027)	.1992(.027)[.027]	.1992(.027)[.027]
ρ	.1621(.100)	.1961(.097)[.099]	.1961(.097)[.098]	.1609(.064)	.1961(.063)[.063]	.1961(.063)[.063]
σ_v^2	.7625(.092)	.9951(.120)[.118]	–	.8782(.073)	.9975(.083)[.082]	–
			$n = 200$; error = 1, 2, 3, for the three panels below			
β_1	1.0002(.019)	1.0001(.019)[.019]	1.0001(.019)[.019]	1.0004(.013)	1.0001(.013)[.013]	1.0001(.013)[.013]
λ	.1964(.028)	.1998(.028)[.029]	.1998(.028)[.029]	.1961(.018)	.1994(.018)[.019]	.1994(.018)[.019]
ρ	.1805(.071)	.1947(.069)[.068]	.1948(.069)[.068]	.1823(.044)	.1985(.044)[.044]	.1986(.044)[.044]
σ_v^2	.7703(.042)	.9958(.054)[.053]	–	.8826(.030)	.9973(.034)[.035]	–
β_1	.9997(.019)	.9996(.019)[.019]	.9996(.019)[.019]	1.0002(.013)	.9999(.013)[.013]	.9999(.013)[.013]
λ	.1969(.029)	.2003(.029)[.029]	.2003(.029)[.028]	.1960(.018)	.1993(.018)[.019]	.1993(.018)[.019]
ρ	.1850(.069)	.1991(.067)[.068]	.1991(.067)[.067]	.1821(.043)	.1984(.043)[.044]	.1984(.043)[.044]
σ_v^2	.7679(.089)	.9927(.115)[.114]	–	.8820(.071)	.9967(.080)[.080]	–
β_1	1.0007(.019)	1.0006(.019)[.019]	1.0006(.019)[.019]	.9996(.012)	.9993(.012)[.013]	.9993(.012)[.013]
λ	.1968(.028)	.2002(.028)[.029]	.2002(.028)[.029]	.1968(.019)	.2000(.019)[.019]	.2000(.019)[.019]
ρ	.1840(.069)	.1981(.067)[.068]	.1981(.067)[.067]	.1818(.044)	.1980(.043)[.044]	.1980(.043)[.044]
σ_v^2	.7688(.063)	.9939(.082)[.083]	–	.8842(.053)	.9992(.060)[.058]	–
			$n = 400$; error = 1, 2, 3, for the three panels below			
β_1	1.0003(.014)	1.0003(.014)[.013]	1.0003(.014)[.013]	1.0004(.009)	1.0003(.009)[.009]	1.0003(.009)[.009]
λ	.1985(.019)	.2001(.019)[.019]	.2001(.019)[.019]	.1982(.014)	.1998(.014)[.014]	.1999(.014)[.014]
ρ	.1936(.049)	.1982(.048)[.047]	.1982(.048)[.047]	.1918(.033)	.1989(.032)[.031]	.1989(.032)[.031]
σ_v^2	.7738(.028)	.9966(.036)[.038]	–	.8854(.022)	.9982(.024)[.025]	–
β_1	1.0001(.013)	1.0000(.013)[.013]	1.0000(.013)[.013]	.9998(.009)	.9997(.009)[.009]	.9997(.009)[.009]
λ	.1985(.019)	.2001(.019)[.020]	.2001(.019)[.019]	.1983(.013)	.1999(.013)[.014]	.2000(.013)[.014]
ρ	.1937(.048)	.1983(.047)[.048]	.1983(.047)[.047]	.1931(.031)	.2001(.030)[.031]	.2001(.030)[.031]
σ_v^2	.7782(.063)	1.0023(.081)[.082]	–	.8847(.050)	.9974(.056)[.057]	–
β_1	1.0001(.013)	1.0001(.013)[.013]	1.0001(.013)[.013]	.9995(.009)	.9994(.009)[.009]	.9994(.009)[.009]
λ	.1972(.020)	.1988(.020)[.020]	.1987(.020)[.019]	.1978(.013)	.1994(.013)[.014]	.1996(.013)[.014]
ρ	.1944(.050)	.1990(.049)[.047]	.1990(.049)[.047]	.1931(.031)	.2002(.031)[.031]	.2002(.031)[.031]
σ_v^2	.7743(.049)	.9973(.063)[.060]	–	.8873(.038)	1.0004(.043)[.042]	–

Note: error = 1(normal), 2(normal mixture), 3(chi-square).The following tables also apply.

Table 1b: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP1` with **homoskedasticity**, Unbalancedness percentage = 10%, $(\beta, \lambda, \rho, \sigma_v^2) = (1, 0.2, 0.2, 1)$, and `W = Group-I` and `M = Queen`.

	T=5			T=10		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
	$n = 50$; error = 1, 2, 3, for the three panels below					
β_1	.9976(.038)	.9986(.038)[.038]	.9986(.038)[.038]	.9993(.025)	.9989(.025)[.026]	.9989(.025)[.026]
λ	.1666(.077)	.1885(.075)[.075]	.1887(.076)[.074]	.1780(.046)	.1971(.045)[.045]	.1971(.046)[.045]
ρ	.1101(.147)	.1889(.141)[.150]	.1889(.141)[.147]	.1210(.093)	.1927(.092)[.091]	.1927(.092)[.091]
σ_v^2	.7390(.082)	.9828(.109)[.106]	–	.8641(.063)	.9936(.072)[.071]	–
β_1	.9980(.038)	.9989(.038)[.038]	.9989(.038)[.038]	.9986(.025)	.9982(.025)[.026]	.9982(.025)[.025]
λ	.1689(.076)	.1909(.074)[.074]	.1909(.074)[.072]	.1779(.047)	.1968(.046)[.045]	.1969(.047)[.045]
ρ	.1121(.148)	.1915(.143)[.150]	.1913(.143)[.143]	.1235(.093)	.1951(.091)[.091]	.1950(.091)[.089]
σ_v^2	.7420(.176)	.9867(.234)[.218]	–	.8641(.138)	.9935(.158)[.156]	–
β_1	.9980(.038)	.9990(.037)[.038]	.9990(.037)[.038]	1.0003(.026)	.9999(.026)[.026]	.9999(.026)[.026]
λ	.1688(.078)	.1907(.076)[.074]	.1908(.076)[.073]	.1771(.046)	.1960(.045)[.045]	.1961(.045)[.045]
ρ	.1104(.150)	.1894(.145)[.150]	.1894(.145)[.146]	.1211(.093)	.1928(.091)[.091]	.1928(.092)[.090]
σ_v^2	.7380(.129)	.9814(.171)[.161]	–	.8615(.101)	.9906(.116)[.113]	–
	$n = 100$; error = 1, 2, 3, for the three panels below					
β_1	1.0015(.029)	1.0009(.029)[.029]	1.0009(.029)[.028]	1.0004(.018)	1.0004(.018)[.017]	1.0004(.018)[.017]
λ	.1842(.055)	.1960(.055)[.054]	.1961(.055)[.053]	.1817(.040)	.1963(.039)[.039]	.1964(.039)[.039]
ρ	.1626(.104)	.1954(.101)[.099]	.1954(.101)[.098]	.1638(.064)	.1988(.063)[.063]	.1989(.063)[.063]
σ_v^2	.7604(.058)	.9928(.076)[.076]	–	.8787(.046)	.9981(.052)[.050]	–
β_1	1.0015(.029)	1.0009(.029)[.029]	1.0009(.029)[.028]	1.0001(.018)	1.0000(.018)[.017]	1.0000(.018)[.017]
λ	.1829(.055)	.1948(.054)[.054]	.1948(.054)[.053]	.1838(.040)	.1983(.040)[.039]	.1983(.040)[.039]
ρ	.1588(.100)	.1917(.097)[.099]	.1916(.097)[.097]	.1601(.063)	.1952(.062)[.063]	.1952(.062)[.062]
σ_v^2	.7674(.128)	1.0019(.167)[.161]	–	.8780(.101)	.9973(.115)[.113]	–
β_1	1.0000(.028)	.9994(.028)[.029]	.9994(.028)[.028]	.9998(.018)	.9998(.018)[.017]	.9998(.018)[.017]
λ	.1831(.056)	.1950(.055)[.054]	.1950(.055)[.053]	.1834(.041)	.1979(.040)[.039]	.1979(.040)[.039]
ρ	.1599(.098)	.1928(.095)[.099]	.1929(.096)[.097]	.1618(.064)	.1969(.063)[.063]	.1969(.063)[.062]
σ_v^2	.7636(.091)	.9970(.118)[.118]	–	.8763(.072)	.9954(.082)[.082]	–
	$n = 200$; error = 1, 2, 3, for the three panels below					
β_1	1.0002(.020)	1.0001(.020)[.020]	1.0001(.020)[.019]	.9996(.013)	.9996(.013)[.012]	.9996(.013)[.012]
λ	.1856(.049)	.1955(.049)[.048]	.1955(.049)[.048]	.1883(.033)	.1986(.033)[.033]	.1986(.033)[.033]
ρ	.1829(.069)	.1970(.068)[.068]	.1970(.068)[.068]	.1834(.044)	.1997(.043)[.044]	.1997(.043)[.044]
σ_v^2	.7708(.040)	.9966(.052)[.053]	–	.8836(.031)	.9986(.035)[.035]	–
β_1	1.0001(.020)	.9999(.020)[.020]	.9999(.020)[.019]	.9998(.013)	.9997(.013)[.012]	.9997(.013)[.012]
λ	.1851(.049)	.1950(.049)[.048]	.1950(.049)[.048]	.1876(.033)	.1979(.033)[.033]	.1980(.033)[.033]
ρ	.1864(.068)	.2004(.066)[.068]	.2004(.066)[.067]	.1808(.045)	.1972(.044)[.044]	.1972(.044)[.044]
σ_v^2	.7701(.091)	.9956(.118)[.114]	–	.8825(.075)	.9973(.084)[.081]	–
β_1	1.0002(.019)	1.0000(.019)[.020]	1.0000(.019)[.020]	1.0005(.012)	1.0005(.012)[.012]	1.0005(.012)[.012]
λ	.1861(.049)	.1960(.048)[.048]	.1960(.048)[.048]	.1878(.033)	.1981(.033)[.033]	.1982(.033)[.033]
ρ	.1832(.070)	.1973(.068)[.068]	.1973(.068)[.067]	.1834(.046)	.1997(.046)[.044]	.1997(.046)[.044]
σ_v^2	.7736(.066)	1.0002(.085)[.085]	–	.8829(.051)	.9978(.057)[.058]	–
	$n = 400$; error = 1, 2, 3, for the three panels below					
β_1	1.0003(.013)	1.0003(.013)[.013]	1.0003(.013)[.013]	1.0001(.009)	1.0000(.009)[.009]	1.0000(.009)[.009]
λ	.1875(.041)	.1949(.040)[.042]	.1949(.040)[.042]	.1922(.027)	.1987(.027)[.026]	.1989(.027)[.026]
ρ	.1953(.049)	.1999(.048)[.047]	.1999(.048)[.047]	.1915(.033)	.1986(.033)[.031]	.1986(.033)[.031]
σ_v^2	.7734(.029)	.9961(.037)[.038]	–	.8853(.022)	.9982(.024)[.025]	–
β_1	1.0007(.013)	1.0007(.013)[.013]	1.0007(.013)[.013]	1.0001(.009)	1.0000(.009)[.009]	1.0000(.009)[.009]
λ	.1899(.041)	.1972(.041)[.042]	.1972(.041)[.042]	.1913(.027)	.1978(.027)[.026]	.1981(.027)[.026]
ρ	.1922(.048)	.1969(.047)[.047]	.1969(.047)[.047]	.1905(.032)	.1976(.032)[.031]	.1976(.032)[.031]
σ_v^2	.7767(.062)	1.0004(.080)[.082]	–	.8851(.051)	.9979(.057)[.057]	–
β_1	.9999(.013)	.9999(.013)[.013]	.9999(.013)[.013]	.9997(.009)	.9996(.009)[.009]	.9996(.009)[.009]
λ	.1921(.042)	.1994(.041)[.042]	.1994(.041)[.042]	.1926(.026)	.1991(.026)[.026]	.1993(.026)[.026]
ρ	.1924(.049)	.1970(.048)[.047]	.1970(.048)[.047]	.1907(.032)	.1978(.031)[.031]	.1978(.031)[.031]
σ_v^2	.7729(.046)	.9955(.059)[.060]	–	.8881(.036)	1.0013(.041)[.042]	–

Table 1c: Empirical mean(sd)[$\hat{se}$] of the `naïve` estimator: `DGP1` with **homoskedasticity**, Unbalancedness percentage = 10%, $(\beta, \lambda, \rho, \sigma_v^2) = (1, 0.2, 0.2, 1)$, and `W = Rook` and `M = Queen`.

		$n = 50$	$n = 100$	$n = 200$	$n = 400$
			error = 1, 2, 3, for the three panels below		
$T = 5$	β	.9978(.047)[.049]	.9989(.034)[.035]	1.0012(.024)[.025]	1.0083(.017)[.018]
	λ	.0909(.064)[.073]	.1183(.048)[.055]	.0928(.026)[.029]	.0989(.021)[.023]
	ρ	.0128(.121)[.176]	.0132(.097)[.128]	.0238(.063)[.076]	.0247(.048)[.058]
	σ_v^2	1.0183(.145)[.144]	1.0559(.097)[.097]	1.0746(.073)[.073]	1.0787(.051)[.049]
	β	1.0011(.048)[.050]	.9987(.034)[.035]	.9997(.025)[.025]	1.0064(.018)[.018]
	λ	.0869(.069)[.078]	.1199(.050)[.054]	.0944(.026)[.029]	.0999(.022)[.023]
	ρ	.0232(.129)[.185]	.0113(.102)[.129]	.0228(.061)[.077]	.0272(.048)[.058]
	σ_v^2	1.0133(.322)[.285]	1.0480(.213)[.194]	1.0819(.155)[.154]	1.0713(.099)[.103]
	β	.9981(.050)[.050]	.9990(.034)[.035]	.9999(.024)[.025]	1.0060(.018)[.018]
	λ	.0898(.064)[.075]	.1161(.048)[.054]	.0938(.027)[.029]	.1014(.021)[.023]
	ρ	.0215(.119)[.180]	.0178(.100)[.127]	.0220(.062)[.076]	.0232(.046)[.058]
	σ_v^2	1.0317(.236)[.214]	1.0536(.153)[.146]	1.0724(.120)[.112]	1.0768(.079)[.076]
			error = 1, 2, 3, for the three panels below		
$T = 10$	β	1.0025(.039)[.040]	1.0036(.026)[.026]	1.0039(.020)[.020]	1.0040(.014)[.016]
	λ	.0733(.043)[.047]	.0750(.031)[.031]	.0594(.021)[.023]	.0632(.016)[.017]
	ρ	.0363(.094)[.117]	.0291(.064)[.073]	.0251(.045)[.053]	.0145(.031)[.039]
	σ_v^2	1.0656(.119)[.113]	1.0570(.077)[.074]	1.0738(.057)[.059]	1.0835(.047)[.045]
	β	1.0008(.038)[.040]	1.0000(.027)[.026]	1.0040(.019)[.020]	1.0040(.015)[.015]
	λ	.0711(.043)[.047]	.0769(.029)[.031]	.0579(.022)[.023]	.0648(.016)[.017]
	ρ	.0345(.093)[.117]	.0247(.062)[.073]	.0242(.043)[.052]	.0162(.029)[.039]
	σ_v^2	1.0809(.254)[.235]	1.0688(.168)[.161]	1.0811(.128)[.129]	1.0813(.104)[.098]
	β	.9999(.040)[.040]	1.0016(.027)[.026]	1.0038(.019)[.020]	1.0033(.015)[.016]
	λ	.0715(.042)[.047]	.0768(.029)[.031]	.0582(.020)[.023]	.0632(.015)[.017]
	ρ	.0425(.094)[.117]	.0307(.064)[.073]	.0263(.039)[.052]	.0158(.029)[.039]
	σ_v^2	1.0625(.180)[.169]	1.0727(.122)[.118]	1.0809(.098)[.095]	1.0853(.073)[.072]

Table 2a: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP2` with **homoskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \beta_2, \lambda, \sigma_v^2) = (1, 0.5, 0.2, 1)$, and `W = Queen`.

	T=5			**T=10**		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
			$n = 50$; error = 1, 2, 3, for the three panels below			
β_1	1.0056(.041)	.9999(.041)[.041]	.9999(.041)[.041]	1.0072(.026)	1.0014(.026)[.027]	1.0014(.026)[.027]
β_2	.5898(.194)	.5147(.195)[.194]	.5146(.195)[.193]	.5823(.122)	.5074(.123)[.124]	.5073(.123)[.125]
λ	.1276(.125)	.1862(.125)[.123]	.1863(.125)[.122]	.1368(.078)	.1935(.078)[.080]	.1935(.079)[.081]
σ_v^2	.7390(.082)	.9779(.109)[.106]	–	.8604(.060)	.9889(.069)[.070]	–
β_1	1.0081(.040)	1.0024(.040)[.041]	1.0024(.040)[.040]	1.0069(.027)	1.0011(.027)[.027]	1.0011(.027)[.027]
β_2	.5909(.196)	.5158(.197)[.195]	.5156(.197)[.189]	.5850(.122)	.5104(.123)[.124]	.5104(.123)[.124]
λ	.1235(.122)	.1822(.122)[.124]	.1823(.123)[.120]	.1354(.080)	.1918(.080)[.079]	.1919(.080)[.079]
σ_v^2	.7410(.180)	.9806(.238)[.216]	–	.8674(.140)	.9970(.161)[.158]	–
β_1	1.0066(.041)	1.0009(.041)[.041]	1.0009(.041)[.041]	1.0068(.027)	1.0010(.027)[.027]	1.0010(.027)[.027]
β_2	.5904(.192)	.5151(.193)[.195]	.5151(.194)[.192]	.5827(.124)	.5082(.125)[.124]	.5079(.125)[.125]
λ	.1252(.123)	.1840(.123)[.124]	.1841(.123)[.122]	.1364(.080)	.1929(.080)[.079]	.1931(.080)[.080]
σ_v^2	.7436(.128)	.9841(.169)[.160]	–	.8607(.105)	.9893(.120)[.113]	–
			$n = 100$; error = 1, 2, 3, for the three panels below			
β_1	1.0047(.030)	1.0016(.030)[.031]	1.0016(.030)[.031]	1.0024(.019)	.9997(.019)[.018]	.9997(.019)[.018]
β_2	.5502(.135)	.5114(.135)[.133]	.5114(.135)[.138]	.5381(.083)	.5025(.083)[.083]	.5025(.083)[.086]
λ	.1621(.085)	.1908(.085)[.084]	.1908(.085)[.088]	.1697(.055)	.1982(.055)[.055]	.1982(.055)[.057]
σ_v^2	.7618(.059)	.9903(.076)[.075]	–	.8765(.043)	.9956(.049)[.050]	–
β_1	1.0053(.031)	1.0021(.031)[.031]	1.0021(.031)[.031]	1.0036(.018)	1.0008(.018)[.018]	1.0008(.018)[.018]
β_2	.5551(.130)	.5163(.130)[.133]	.5162(.130)[.136]	.5403(.083)	.5046(.083)[.083]	.5045(.083)[.086]
λ	.1585(.084)	.1872(.084)[.084]	.1872(.084)[.087]	.1678(.056)	.1963(.056)[.055]	.1963(.056)[.057]
σ_v^2	.7675(.129)	.9977(.168)[.159]	–	.8771(.101)	.9963(.115)[.113]	–
β_1	1.0032(.030)	1.0001(.030)[.030]	1.0001(.030)[.031]	1.0029(.018)	1.0001(.018)[.018]	1.0001(.018)[.018]
β_2	.5535(.136)	.5149(.136)[.133]	.5150(.136)[.136]	.5381(.083)	.5024(.083)[.083]	.5024(.083)[.086]
λ	.1598(.086)	.1884(.085)[.084]	.1883(.085)[.087]	.1686(.055)	.1971(.055)[.055]	.1971(.055)[.057]
σ_v^2	.7616(.091)	.9900(.119)[.116]	–	.8755(.074)	.9944(.084)[.081]	–
			$n = 200$; error = 1, 2, 3, for the three panels below			
β_1	1.0020(.021)	1.0006(.021)[.020]	1.0006(.021)[.021]	1.0017(.013)	1.0003(.013)[.013]	1.0003(.013)[.013]
β_2	.5244(.096)	.5056(.096)[.094]	.5056(.096)[.097]	.5218(.061)	.5034(.061)[.060]	.5034(.061)[.062]
λ	.1824(.058)	.1962(.059)[.057]	.1962(.059)[.060]	.1827(.039)	.1972(.039)[.039]	.1972(.039)[.040]
σ_v^2	.7713(.041)	.9948(.053)[.053]	–	.8835(.032)	.9979(.036)[.035]	–
β_1	1.0013(.020)	.9999(.020)[.020]	.9999(.020)[.021]	1.0018(.013)	1.0005(.013)[.013]	1.0005(.013)[.013]
β_2	.5269(.093)	.5082(.093)[.093]	.5081(.093)[.097]	.5202(.060)	.5017(.060)[.060]	.5018(.060)[.062]
λ	.1808(.057)	.1945(.057)[.057]	.1946(.057)[.060]	.1838(.040)	.1982(.040)[.039]	.1982(.040)[.040]
σ_v^2	.7726(.091)	.9964(.118)[.114]	–	.8816(.073)	.9958(.082)[.080]	–
β_1	1.0017(.020)	1.0002(.020)[.020]	1.0002(.020)[.021]	1.0017(.013)	1.0003(.013)[.013]	1.0003(.013)[.013]
β_2	.5248(.094)	.5060(.094)[.094]	.5060(.094)[.097]	.5220(.060)	.5035(.060)[.060]	.5035(.060)[.062]
λ	.1826(.056)	.1963(.056)[.057]	.1963(.056)[.060]	.1834(.039)	.1978(.039)[.039]	.1978(.039)[.040]
σ_v^2	.7741(.065)	.9984(.083)[.084]	–	.8837(.052)	.9981(.059)[.058]	–
			$n = 400$; error = 1, 2, 3, for the three panels below			
β_1	1.0012(.014)	1.0005(.014)[.014]	1.0005(.014)[.014]	1.0007(.009)	1.0000(.009)[.009]	1.0000(.009)[.010]
β_2	.5120(.065)	.5028(.065)[.064]	.5028(.065)[.067]	.5101(.043)	.5004(.043)[.044]	.5004(.043)[.046]
λ	.1909(.041)	.1980(.041)[.041]	.1980(.041)[.043]	.1918(.026)	.1990(.026)[.028]	.1989(.026)[.029]
σ_v^2	.7761(.030)	.9992(.039)[.038]	–	.8851(.022)	.9980(.025)[.025]	–
β_1	1.0011(.014)	1.0004(.014)[.014]	1.0004(.014)[.014]	1.0006(.009)	.9998(.009)[.009]	.9998(.009)[.010]
β_2	.5116(.064)	.5024(.064)[.064]	.5024(.064)[.067]	.5086(.043)	.4989(.043)[.044]	.5018(.060)[.062]
λ	.1904(.042)	.1975(.042)[.041]	.1975(.042)[.043]	.1934(.026)	.2006(.026)[.028]	.1982(.040)[.040]
σ_v^2	.7730(.062)	.9952(.080)[.081]	–	.8867(.052)	.9998(.059)[.058]	–
β_1	1.0010(.014)	1.0003(.014)[.014]	1.0003(.014)[.014]	1.0009(.009)	1.0002(.009)[.009]	1.0003(.013)[.013]
β_2	.5101(.063)	.5009(.063)[.064]	.5009(.063)[.067]	.5127(.042)	.5030(.042)[.044]	.5035(.060)[.062]
λ	.1913(.040)	.1984(.040)[.041]	.1984(.040)[.043]	.1913(.026)	.1984(.026)[.028]	.1978(.039)[.040]
σ_v^2	.7764(.047)	.9996(.061)[.060]	–	.8840(.037)	.9967(.042)[.041]	–

Table 2b: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP2` with **homoskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \beta_2, \lambda, \sigma_v^2) = (1, 0.5, 0.2, 1)$, and `W = Group-I`.

	T=5			**T=10**		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
	$n = 50$; error = 1, 2, 3, for the three panels below					
β_1	1.0130(.043)	1.0060(.043)[.043]	1.0060(.043)[.042]	1.0083(.029)	1.0016(.029)[.028]	1.0016(.029)[.028]
β_2	.6567(.251)	.5636(.241)[.231]	.5637(.242)[.228]	.6090(.159)	.5238(.152)[.149]	.5237(.152)[.149]
λ	.1103(.131)	.1644(.125)[.119]	.1643(.126)[.117]	.1357(.081)	.1861(.078)[.075]	.1861(.078)[.075]
σ_v^2	.7398(.081)	.9793(.108)[.106]	–	.8607(.062)	.9894(.071)[.071]	–
β_1	1.0101(.044)	1.0032(.044)[.042]	1.0031(.044)[.042]	1.0099(.028)	1.0032(.027)[.028]	1.0032(.027)[.028]
β_2	.6426(.238)	.5504(.229)[.229]	.5497(.230)[.220]	.6074(.156)	.5223(.150)[.149]	.5224(.150)[.148]
λ	.1130(.124)	.1668(.119)[.119]	.1673(.120)[.114]	.1355(.079)	.1858(.075)[.075]	.1857(.075)[.073]
σ_v^2	.7413(.171)	.9812(.226)[.215]	–	.8686(.143)	.9984(.164)[.159]	–
β_1	1.0113(.043)	1.0044(.042)[.042]	1.0044(.042)[.042]	1.0069(.028)	1.0003(.028)[.028]	1.0003(.028)[.028]
β_2	.6385(.242)	.5463(.232)[.229]	.5458(.234)[.225]	.6057(.153)	.5208(.147)[.148]	.5209(.147)[.148]
λ	.1183(.126)	.1719(.121)[.118]	.1723(.122)[.116]	.1370(.078)	.1872(.074)[.074]	.1871(.074)[.074]
σ_v^2	.7411(.130)	.9809(.172)[.158]	–	.8575(.101)	.9857(.116)[.112]	–
	$n = 100$; error = 1, 2, 3, for the three panels below					
β_1	1.0038(.027)	1.0012(.027)[.027]	1.0012(.027)[.027]	1.0033(.019)	.9999(.019)[.019]	.9999(.019)[.019]
β_2	.5879(.180)	.5307(.175)[.173]	.5306(.176)[.177]	.5815(.135)	.5142(.130)[.127]	.5140(.131)[.128]
λ	.1466(.095)	.1808(.093)[.092]	.1808(.093)[.094]	.1528(.069)	.1927(.067)[.065]	.1928(.067)[.066]
σ_v^2	.7630(.058)	.9921(.075)[.075]	–	.8766(.044)	.9958(.050)[.050]	–
β_1	1.0043(.027)	1.0016(.027)[.027]	1.0016(.027)[.027]	1.0036(.019)	1.0002(.019)[.019]	1.0002(.019)[.019]
β_2	.5956(.189)	.5384(.183)[.173]	.5385(.183)[.175]	.5882(.133)	.5207(.128)[.128]	.5206(.128)[.128]
λ	.1433(.100)	.1775(.097)[.092]	.1774(.097)[.093]	.1480(.069)	.1881(.066)[.066]	.1882(.066)[.066]
σ_v^2	.7644(.129)	.9940(.167)[.159]	–	.8752(.101)	.9942(.114)[.113]	–
β_1	1.0044(.027)	1.0017(.027)[.027]	1.0017(.027)[.027]	1.0045(.019)	1.0011(.019)[.019]	1.0011(.019)[.019]
β_2	.5859(.180)	.5285(.175)[.173]	.5283(.175)[.176]	.5853(.131)	.5179(.127)[.127]	.5179(.126)[.128]
λ	.1465(.096)	.1808(.093)[.092]	.1810(.093)[.094]	.1488(.068)	.1889(.066)[.066]	.1889(.066)[.066]
σ_v^2	.7676(.095)	.9981(.123)[.118]	–	.8755(.075)	.9945(.086)[.081]	–
	$n = 200$; error = 1, 2, 3, for the three panels below					
β_1	1.0027(.020)	1.0011(.020)[.020]	1.0011(.020)[.020]	1.0023(.013)	1.0007(.013)[.013]	1.0007(.013)[.013]
β_2	.5722(.170)	.5257(.165)[.157]	.5257(.165)[.160]	.5592(.109)	.5124(.106)[.104]	.5124(.106)[.106]
λ	.1597(.083)	.1858(.081)[.079]	.1859(.081)[.080]	.1653(.055)	.1926(.054)[.054]	.1925(.054)[.055]
σ_v^2	.7726(.041)	.9949(.053)[.053]	–	.8837(.032)	.9983(.036)[.035]	–
β_1	1.0026(.020)	1.0011(.020)[.020]	1.0011(.020)[.020]	1.0021(.013)	1.0005(.013)[.013]	1.0005(.013)[.013]
β_2	.5687(.160)	.5224(.156)[.157]	.5222(.156)[.158]	.5580(.106)	.5112(.103)[.104]	.5110(.103)[.106]
λ	.1605(.080)	.1866(.078)[.079]	.1867(.078)[.079]	.1661(.055)	.1933(.053)[.054]	.1934(.053)[.055]
σ_v^2	.7740(.089)	.9967(.114)[.113]	–	.8839(.072)	.9985(.082)[.081]	–
β_1	1.0033(.019)	1.0018(.019)[.020]	1.0018(.019)[.020]	1.0019(.013)	1.0004(.013)[.013]	1.0004(.013)[.013]
β_2	.5759(.164)	.5293(.160)[.158]	.5293(.160)[.159]	.5581(.107)	.5113(.104)[.104]	.5112(.104)[.106]
λ	.1564(.082)	.1827(.080)[.079]	.1827(.080)[.080]	.1670(.054)	.1942(.053)[.054]	.1943(.053)[.055]
σ_v^2	.7726(.067)	.9950(.087)[.083]	–	.8844(.051)	.9991(.057)[.058]	–
	$n = 400$; error = 1, 2, 3, for the three panels below					
β_1	1.0012(.014)	1.0005(.014)[.014]	1.0005(.014)[.014]	1.0012(.009)	1.0004(.009)[.009]	1.0004(.009)[.009]
β_2	.5526(.142)	.5197(.139)[.136]	.5196(.139)[.138]	.5476(.100)	.5120(.098)[.091]	.5121(.098)[.093]
λ	.1702(.069)	.1890(.067)[.067]	.1891(.067)[.068]	.1732(.049)	.1935(.048)[.046]	.1934(.048)[.047]
σ_v^2	.7759(.029)	.9989(.038)[.038]	–	.8867(.023)	.9998(.026)[.025]	–
β_1	1.0010(.014)	1.0003(.014)[.014]	1.0003(.014)[.014]	1.0007(.009)	.9999(.009)[.009]	.9999(.009)[.009]
β_2	.5553(.141)	.5224(.139)[.136]	.5225(.139)[.138]	.5444(.093)	.5088(.091)[.091]	.5088(.091)[.092]
λ	.1682(.070)	.1870(.069)[.067]	.1870(.069)[.068]	.1753(.047)	.1955(.046)[.046]	.1956(.046)[.047]
σ_v^2	.7756(.064)	.9986(.082)[.082]	–	.8853(.051)	.9983(.058)[.057]	–
β_1	1.0012(.014)	1.0004(.014)[.014]	1.0004(.014)[.014]	1.0012(.009)	1.0003(.009)[.009]	1.0003(.009)[.009]
β_2	.5551(.140)	.5223(.138)[.136]	.5222(.138)[.138]	.5489(.094)	.5131(.092)[.092]	.5131(.092)[.093]
λ	.1692(.068)	.1879(.066)[.067]	.1880(.067)[.068]	.1726(.048)	.1930(.047)[.046]	.1930(.047)[.047]
σ_v^2	.7763(.048)	.9996(.061)[.060]	–	.8864(.036)	.9996(.041)[.042]	–

Table 3a: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP1` with **heteroskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \lambda, \rho, \sigma_v^2) = (1, 0.2, 0.2, 1)$, and `W = M = Group-II`.

	T=5			**T=10**		
	QMLE	M-Est	M-Est	QMLE	M-Est	RM-Est
			$n = 50$; error = 1, 2, 3, for the three panels below			
β_1	1.0001(.042)	1.0001(.042)[.041]	.9993(.042)[.042]	1.0015(.025)	1.0012(.025)[.025]	1.0009(.025)[.024]
λ	.1878(.070)	.1944(.068)[.098]	.1973(.080)[.080]	.1899(.039)	.1953(.038)[.054]	.1987(.043)[.042]
ρ	-.0161(.209)	.0905(.177)[.199]	.1016(.272)[.247]	.0447(.119)	.1342(.106)[.117]	.1660(.146)[.139]
σ_v^2	.7664(.102)	1.0237(.136)[.150]	–	.8647(.073)	.9940(.083)[.093]	–
β_1	1.0012(.042)	1.0013(.042)[.040]	1.0005(.043)[.041]	.9997(.025)	.9995(.025)[.025]	.9992(.025)[.024]
λ	.1873(.072)	.1938(.069)[.098]	.1956(.082)[.080]	.1900(.039)	.1954(.037)[.054]	.1990(.042)[.042]
ρ	-.0008(.198)	.1036(.168)[.198]	.1235(.248)[.232]	.0425(.118)	.1324(.105)[.119]	.1636(.145)[.137]
σ_v^2	.7606(.217)	1.0154(.290)[.274]	–	.8715(.175)	1.0019(.201)[.192]	–
β_1	.9986(.041)	.9986(.042)[.041]	.9978(.042)[.041]	1.0011(.025)	1.0009(.025)[.025]	1.0006(.025)[.024]
λ	.1843(.070)	.1911(.067)[.099]	.1928(.081)[.080]	.1898(.039)	.1952(.037)[.054]	.1987(.042)[.042]
ρ	-.0072(.205)	.0980(.174)[.198]	.1144(.260)[.238]	.0461(.115)	.1355(.102)[.117]	.1682(.139)[.138]
σ_v^2	.7727(.158)	1.0319(.211)[.212]	–	.8646(.125)	.9940(.144)[.141]	–
			$n = 100$; error = 1, 2, 3, for the three panels below			
β_1	1.0005(.028)	1.0005(.028)[.028]	1.0002(.028)[.028]	.9994(.018)	.9995(.018)[.018]	.9996(.018)[.018]
λ	.1927(.049)	.1954(.048)[.064]	.1992(.056)[.054]	.1901(.034)	.1936(.034)[.043]	.1980(.040)[.038]
ρ	.0883(.131)	.1304(.120)[.127]	.1573(.167)[.160]	.1077(.082)	.1470(.077)[.080]	.1809(.106)[.101]
σ_v^2	.7572(.071)	.9925(.093)[.105]	–	.8663(.053)	.9860(.060)[.067]	–
β_1	1.0004(.029)	1.0004(.029)[.028]	1.0001(.029)[.028]	.9991(.018)	.9992(.018)[.018]	.9992(.018)[.018]
λ	.1921(.049)	.1948(.048)[.063]	.1985(.055)[.053]	.1902(.033)	.1937(.032)[.043]	.1978(.038)[.038]
ρ	.0884(.129)	.1305(.118)[.128]	.1578(.165)[.157]	.1106(.078)	.1497(.073)[.080]	.1848(.101)[.099]
σ_v^2	.7554(.155)	.9901(.203)[.199]	–	.8659(.124)	.9855(.141)[.139]	–
β_1	.9997(.029)	.9997(.029)[.028]	.9995(.029)[.028]	.9999(.018)	1.0000(.018)[.018]	1.0000(.018)[.018]
λ	.1914(.049)	.1941(.048)[.063]	.1979(.055)[.054]	.1922(.033)	.1957(.032)[.043]	.2005(.037)[.038]
ρ	.0877(.130)	.1299(.119)[.128]	.1566(.167)[.159]	.1077(.079)	.1470(.074)[.080]	.1808(.102)[.100]
σ_v^2	.7614(.115)	.9979(.150)[.152]	–	.8630(.088)	.9822(.100)[.102]	–
			$n = 200$; error = 1, 2, 3, for the three panels below			
β_1	.9991(.019)	.9991(.019)[.019]	.9990(.019)[.019]	.9999(.013)	1.0000(.013)[.013]	1.0000(.013)[.013]
λ	.1950(.034)	.1962(.034)[.044]	.2000(.040)[.040]	.1938(.026)	.1951(.026)[.030]	.1976(.030)[.029]
ρ	.1272(.086)	.1441(.082)[.084]	.1763(.112)[.107]	.1420(.057)	.1601(.055)[.055]	.1955(.071)[.069]
σ_v^2	.7657(.050)	.9907(.065)[.073]	–	.8932(.038)	1.0097(.043)[.048]	–
β_1	.9996(.019)	.9996(.019)[.019]	.9995(.019)[.019]	.9999(.013)	1.0000(.013)[.013]	1.0000(.013)[.013]
λ	.1939(.035)	.1951(.035)[.044]	.1984(.041)[.040]	.1958(.026)	.1972(.025)[.030]	.2000(.029)[.029]
ρ	.1345(.083)	.1511(.079)[.083]	.1857(.106)[.105]	.1399(.056)	.1581(.054)[.055]	.1926(.070)[.069]
σ_v^2	.7704(.111)	.9967(.144)[.143]	–	.8939(.089)	1.0105(.101)[.101]	–
β_1	.9996(.019)	.9996(.019)[.019]	.9996(.019)[.019]	.9993(.013)	.9993(.013)[.013]	.9993(.013)[.013]
λ	.1947(.035)	.1959(.035)[.044]	.1996(.041)[.040]	.1955(.026)	.1969(.026)[.030]	.1998(.029)[.029]
ρ	.1290(.085)	.1459(.081)[.084]	.1787(.110)[.106]	.1379(.057)	.1561(.055)[.055]	.1903(.071)[.070]
σ_v^2	.7646(.080)	.9893(.104)[.106]	–	.8916(.063)	1.0080(.072)[.074]	–
			$n = 400$; error = 1, 2, 3, for the three panels below			
β_1	.9999(.014)	.9999(.014)[.013]	.9999(.014)[.014]	1.0003(.009)	1.0009(.009)[.009]	1.0003(.009)[.009]
λ	.1966(.026)	.1970(.026)[.031]	.1998(.030)[.030]	.1974(.016)	.2490(.018)[.019]	.2010(.018)[.018]
ρ	.1491(.060)	.1550(.058)[.058]	.1892(.075)[.074]	.1533(.037)	.1210(.028)[.037]	.1959(.046)[.047]
σ_v^2	.7849(.034)	1.0110(.044)[.052]	–	.8923(.027)	1.0063(.030)[.034]	–
β_1	.9998(.014)	.9998(.014)[.013]	.9998(.014)[.014]	.9995(.009)	1.0001(.009)[.009]	.9995(.009)[.009]
λ	.1968(.027)	.1972(.027)[.031]	.1998(.031)[.030]	.1965(.017)	.2493(.019)[.019]	.1997(.019)[.018]
ρ	.1509(.061)	.1568(.059)[.058]	.1914(.075)[.074]	.1562(.038)	.1232(.029)[.037]	.1996(.048)[.047]
σ_v^2	.7878(.079)	1.0148(.102)[.103]	–	.8933(.061)	1.0075(.069)[.072]	–
β_1	1.0000(.013)	1.0000(.013)[.013]	1.0000(.014)[.014]	.9998(.009)	1.0004(.009)[.009]	.9998(.009)[.009]
λ	.1949(.027)	.1953(.027)[.031]	.1980(.031)[.030]	.1968(.017)	.2485(.018)[.019]	.2003(.019)[.018]
ρ	.1500(.059)	.1559(.057)[.058]	.1904(.073)[.074]	.1531(.038)	.1208(.029)[.037]	.1958(.047)[.047]
σ_v^2	.7869(.059)	1.0136(.076)[.078]	–	.8960(.045)	1.0105(.051)[.053]	–

Table 3b: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP1` with **heteroskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \lambda, \rho, \sigma_v^2) = (1, 0.2, 0.2, 1)$, and `W = Group-II, M = Queen`.

	T=5			T=10		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
			$n = 50$; error = 1, 2, 3, for the three panels below			
β_1	.9974(.039)	.9978(.039)[.040]	.9979(.039)[.039]	1.0024(.026)	1.0007(.026)[.026]	.9999(.026)[.025]
λ	.1311(.086)	.1735(.083)[.106]	.1894(.089)[.090]	.1677(.036)	.1880(.035)[.047]	.1973(.037)[.037]
ρ	.1077(.147)	.1845(.141)[.145]	.1889(.141)[.142]	.1290(.092)	.2001(.091)[.090]	.1988(.091)[.089]
σ_v^2	.7717(.102)	1.0264(.136)[.150]	–	.8737(.074)	1.0030(.085)[.094]	–
β_1	.9980(.039)	.9985(.039)[.040]	.9985(.039)[.038]	1.0015(.025)	.9998(.025)[.025]	.9989(.025)[.025]
λ	.1343(.083)	.1757(.080)[.105]	.1915(.085)[.088]	.1686(.035)	.1888(.034)[.047]	.1981(.036)[.037]
ρ	.1086(.146)	.1853(.140)[.146]	.1894(.140)[.135]	.1284(.092)	.1995(.090)[.090]	.1980(.089)[.087]
σ_v^2	.7695(.221)	1.0234(.294)[.277]	–	.8744(.173)	1.0037(.198)[.193]	–
β_1	.9980(.040)	.9984(.040)[.040]	.9985(.040)[.039]	1.0021(.025)	1.0004(.025)[.026]	.9995(.025)[.025]
λ	.1316(.084)	.1737(.081)[.106]	.1896(.087)[.089]	.1677(.035)	.1881(.035)[.047]	.1974(.036)[.037]
ρ	.1113(.144)	.1878(.138)[.145]	.1919(.138)[.139]	.1280(.092)	.1992(.090)[.090]	.1981(.090)[.088]
σ_v^2	.7744(.161)	1.0300(.214)[.210]	–	.8751(.123)	1.0047(.141)[.144]	–
			$n = 100$; error = 1, 2, 3, for the three panels below			
β_1	1.0003(.025)	1.0005(.025)[.026]	1.0006(.025)[.026]	.9996(.018)	.9994(.018)[.018]	.9992(.018)[.018]
λ	.1733(.045)	.1849(.045)[.054]	.1960(.046)[.048]	.1760(.034)	.1889(.033)[.038]	.2000(.035)[.035]
ρ	.1665(.095)	.1995(.093)[.095]	.1980(.092)[.094]	.1627(.066)	.1979(.065)[.064]	.1979(.065)[.064]
σ_v^2	.7706(.070)	.9985(.090)[.102]	–	.8717(.054)	.9931(.062)[.067]	–
β_1	.9994(.026)	.9996(.026)[.026]	.9997(.026)[.026]	1.0003(.019)	1.0001(.019)[.018]	.9999(.019)[.018]
λ	.1736(.045)	.1851(.045)[.054]	.1960(.046)[.047]	.1746(.034)	.1875(.033)[.038]	.1985(.035)[.035]
ρ	.1668(.093)	.1997(.090)[.095]	.1984(.090)[.090]	.1605(.066)	.1957(.065)[.064]	.1957(.065)[.063]
σ_v^2	.7648(.153)	.9910(.198)[.194]	–	.8736(.125)	.9953(.143)[.139]	–
β_1	.9996(.026)	.9998(.026)[.026]	.9999(.026)[.026]	.9997(.019)	.9995(.019)[.018]	.9993(.019)[.018]
λ	.1739(.045)	.1854(.045)[.054]	.1964(.047)[.047]	.1738(.033)	.1867(.033)[.038]	.1977(.035)[.035]
ρ	.1656(.097)	.1985(.095)[.095]	.1971(.094)[.093]	.1607(.066)	.1958(.065)[.064]	.1958(.064)[.063]
σ_v^2	.7646(.112)	.9907(.145)[.146]	–	.8760(.091)	.9981(.104)[.103]	–
			$n = 200$; error = 1, 2, 3, for the three panels below			
β_1	.9987(.020)	.9988(.020)[.019]	.9989(.020)[.019]	1.0004(.013)	1.0001(.013)[.013]	.9997(.013)[.013]
λ	.1784(.035)	.1853(.034)[.042]	.1980(.036)[.037]	.1833(.022)	.1892(.022)[.026]	.1997(.023)[.023]
ρ	.1862(.071)	.2006(.070)[.068]	.1996(.069)[.069]	.1830(.045)	.1995(.044)[.044]	.1999(.044)[.044]
σ_v^2	.7647(.051)	.9883(.065)[.073]	–	.8922(.036)	1.0082(.041)[.048]	–
β_1	.9992(.019)	.9993(.019)[.019]	.9994(.019)[.019]	1.0007(.012)	1.0005(.012)[.013]	1.0000(.012)[.013]
λ	.1773(.035)	.1842(.034)[.042]	.1968(.036)[.037]	.1821(.022)	.1880(.022)[.026]	.1985(.023)[.023]
ρ	.1816(.071)	.1960(.069)[.068]	.1952(.069)[.068]	.1798(.045)	.1964(.045)[.044]	.1968(.045)[.044]
σ_v^2	.7660(.110)	.9899(.142)[.141]	–	.8912(.090)	1.0070(.102)[.100]	–
β_1	.9987(.019)	.9988(.019)[.019]	.9990(.019)[.019]	1.0006(.013)	1.0003(.013)[.013]	.9999(.013)[.013]
λ	.1782(.035)	.1851(.034)[.042]	.1978(.036)[.037]	.1830(.023)	.1890(.023)[.026]	.1996(.024)[.023]
ρ	.1840(.070)	.1984(.069)[.068]	.1975(.068)[.068]	.1784(.043)	.1949(.043)[.044]	.1953(.043)[.044]
σ_v^2	.7664(.081)	.9905(.104)[.106]	–	.8956(.065)	1.0121(.074)[.074]	–
			$n = 400$; error = 1, 2, 3, for the three panels below			
β_1	1.0000(.013)	1.0000(.013)[.013]	1.0000(.013)[.013]	.9997(.009)	.9997(.009)[.009]	.9994(.009)[.009]
λ	.1862(.024)	.1894(.024)[.028]	.2002(.025)[.026]	.1860(.016)	.1888(.016)[.018]	.1991(.016)[.017]
ρ	.1945(.047)	.1990(.046)[.047]	.1991(.046)[.047]	.1924(.032)	.1995(.031)[.031]	.2001(.031)[.032]
σ_v^2	.7839(.034)	1.0096(.044)[.052]	–	.8913(.028)	1.0049(.031)[.034]	–
β_1	.9994(.013)	.9994(.013)[.013]	.9994(.013)[.013]	1.0003(.009)	1.0003(.009)[.009]	1.0000(.009)[.009]
λ	.1835(.025)	.1866(.025)[.028]	.1973(.026)[.026]	.1871(.016)	.1900(.016)[.018]	.2004(.017)[.017]
ρ	.1959(.048)	.2003(.047)[.047]	.2004(.047)[.047]	.1914(.031)	.1985(.030)[.031]	.1991(.030)[.031]
σ_v^2	.7842(.080)	1.0100(.103)[.103]	–	.8923(.063)	1.0060(.071)[.071]	–
β_1	1.0005(.014)	1.0005(.014)[.013]	1.0005(.014)[.013]	1.0003(.009)	1.0002(.009)[.009]	1.0000(.009)[.009]
λ	.1861(.024)	.1893(.024)[.028]	.2001(.025)[.026]	.1864(.016)	.1892(.016)[.018]	.1996(.016)[.017]
ρ	.1955(.047)	.1999(.046)[.047]	.2001(.046)[.047]	.1922(.031)	.1993(.031)[.031]	.1999(.031)[.032]
σ_v^2	.7854(.059)	1.0116(.076)[.078]	–	.8950(.047)	1.0090(.053)[.053]	–

Table 4: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP2` with **heteroskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \beta_2, \lambda, \sigma_v^2) = (1, 0.5, 0.2, 1)$, and `W = Group-II`.

	T=5			**T=10**		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
			$n = 50$; error = 1, 2, 3, for the three panels below			
β_1	1.0113(.045)	1.0054(.045)[.042]	1.0033(.045)[.044]	1.0134(.028)	1.0058(.028)[.027]	1.0024(.028)[.029]
β_2	.7304(.237)	.6103(.223)[.277]	.5672(.256)[.269]	.6962(.153)	.5798(.143)[.172]	.5278(.167)[.190]
λ	.0569(.146)	.1317(.137)[.152]	.1583(.159)[.165]	.0811(.089)	.1517(.084)[.096]	.1831(.099)[.114]
σ_v^2	.7704(.102)	1.0208(.135)[.148]	–	.8676(.074)	.9945(.085)[.093]	–
β_1	1.0124(.045)	1.0065(.044)[.042]	1.0045(.045)[.043]	1.0130(.028)	1.0054(.028)[.027]	1.0021(.028)[.028]
β_2	.7306(.230)	.6121(.216)[.276]	.5705(.247)[.261]	.6972(.151)	.5811(.141)[.172]	.5302(.166)[.187]
λ	.0572(.141)	.1310(.132)[.151]	.1566(.153)[.160]	.0805(.089)	.1508(.083)[.096]	.1816(.099)[.112]
σ_v^2	.7664(.224)	1.0156(.297)[.273]	–	.8662(.171)	.9930(.196)[.189]	–
β_1	1.0084(.044)	1.0025(.044)[.043]	1.0003(.044)[.043]	1.0127(.028)	1.0051(.028)[.027]	1.0017(.029)[.029]
β_2	.7193(.232)	.5998(.218)[.276]	.5557(.249)[.266]	.6969(.156)	.5804(.146)[.172]	.5285(.171)[.189]
λ	.0627(.143)	.1373(.134)[.152]	.1645(.155)[.164]	.0809(.092)	.1514(.086)[.096]	.1828(.102)[.113]
σ_v^2	.7771(.166)	1.0297(.219)[.211]	–	.8710(.122)	.9984(.140)[.142]	–
			$n = 100$; error = 1, 2, 3, for the three panels below			
β_1	1.0090(.032)	1.0056(.032)[.030]	1.0028(.032)[.032]	1.0080(.020)	1.0045(.020)[.019]	1.0011(.020)[.021]
β_2	.6379(.166)	.5817(.161)[.185]	.5351(.185)[.201]	.6174(.106)	.5644(.102)[.118]	.5130(.120)[.125]
λ	.1149(.095)	.1497(.092)[.103]	.1785(.109)[.119]	.1250(.063)	.1588(.061)[.066]	.1914(.073)[.077]
σ_v^2	.7609(.070)	.9942(.091)[.105]	–	.8659(.052)	.9843(.060)[.067]	–
β_1	1.0081(.031)	1.0047(.031)[.029]	1.0019(.031)[.032]	1.0080(.020)	1.0045(.020)[.019]	1.0012(.020)[.021]
β_2	.6336(.161)	.5778(.156)[.185]	.5313(.179)[.196]	.6179(.104)	.5650(.101)[.118]	.5141(.118)[.124]
λ	.1168(.092)	.1513(.089)[.102]	.1801(.105)[.116]	.1247(.063)	.1584(.061)[.067]	.1907(.072)[.076]
σ_v^2	.7653(.160)	.9998(.209)[.200]	–	.8667(.119)	.9853(.136)[.138]	–
β_1	1.0081(.031)	1.0048(.031)[.030]	1.0020(.031)[.032]	1.0077(.020)	1.0042(.020)[.019]	1.0009(.021)[.021]
β_2	.6370(.163)	.5810(.158)[.185]	.5348(.182)[.200]	.6186(.106)	.5656(.103)[.118]	.5150(.121)[.124]
λ	.1150(.095)	.1497(.092)[.103]	.1782(.108)[.118]	.1238(.064)	.1575(.062)[.067]	.1897(.075)[.076]
σ_v^2	.7629(.115)	.9968(.150)[.152]	–	.8691(.090)	.9880(.102)[.102]	–
			$n = 200$; error = 1, 2, 3, for the three panels below			
β_1	1.0067(.021)	1.0050(.021)[.021]	1.0017(.022)[.023]	1.0053(.015)	1.0036(.015)[.014]	1.0005(.015)[.015]
β_2	.5960(.117)	.5690(.116)[.129]	.5193(.133)[.146]	.5831(.077)	.5563(.076)[.082]	.5074(.086)[.093]
λ	.1412(.067)	.1573(.066)[.068]	.1869(.077)[.085]	.1485(.044)	.1651(.044)[.046]	.1953(.051)[.055]
σ_v^2	.7670(.051)	.9909(.066)[.073]	–	.8830(.039)	.9979(.044)[.047]	–
β_1	1.0056(.021)	1.0038(.021)[.021]	1.0005(.022)[.023]	1.0051(.014)	1.0034(.014)[.014]	1.0003(.015)[.015]
β_2	.5952(.113)	.5683(.111)[.129]	.5179(.127)[.144]	.5807(.075)	.5540(.073)[.083]	.5048(.083)[.093]
λ	.1438(.064)	.1598(.063)[.068]	.1898(.074)[.084]	.1495(.043)	.1660(.043)[.046]	.1964(.050)[.055]
σ_v^2	.7684(.110)	.9926(.143)[.141]	–	.8814(.088)	.9962(.099)[.099]	–
β_1	1.0062(.022)	1.0044(.021)[.021]	1.0011(.022)[.023]	1.0049(.015)	1.0032(.015)[.014]	1.0001(.015)[.015]
β_2	.5952(.116)	.5682(.114)[.129]	.5180(.131)[.145]	.5839(.077)	.5572(.076)[.083]	.5081(.086)[.093]
λ	.1434(.066)	.1595(.065)[.068]	.1893(.076)[.084]	.1487(.044)	.1652(.043)[.046]	.1956(.051)[.055]
σ_v^2	.7646(.082)	.9878(.105)[.106]	–	.8848(.063)	1.0000(.071)[.074]	–
			$n = 400$; error = 1, 2, 3, for the three panels below			
β_1	1.0049(.015)	1.0040(.015)[.014]	1.0006(.015)[.016]	1.0039(.010)	1.0030(.010)[.010]	1.0000(.010)[.011]
β_2	.5716(.079)	.5582(.078)[.086]	.5083(.090)[.103]	.5631(.051)	.5498(.051)[.057]	.5031(.057)[.065]
λ	.1562(.046)	.1643(.046)[.048]	.1947(.054)[.062]	.1616(.030)	.1698(.030)[.032]	.1987(.034)[.039]
σ_v^2	.7778(.036)	.9995(.046)[.052]	–	.8933(.027)	1.0073(.031)[.034]	–
β_1	1.0048(.014)	1.0040(.014)[.014]	1.0006(.015)[.015]	1.0041(.010)	1.0032(.010)[.010]	1.0003(.010)[.010]
β_2	.5695(.080)	.5561(.079)[.086]	.5057(.091)[.103]	.5627(.051)	.5495(.051)[.057]	.5030(.057)[.065]
λ	.1579(.047)	.1660(.047)[.048]	.1968(.055)[.061]	.1609(.030)	.1691(.029)[.032]	.1979(.034)[.039]
σ_v^2	.7778(.078)	.9995(.100)[.102]	–	.8948(.064)	1.0089(.072)[.072]	–
β_1	1.0047(.014)	1.0038(.014)[.014]	1.0005(.014)[.015]	1.0038(.010)	1.0030(.010)[.010]	1.0000(.010)[.010]
β_2	.5714(.078)	.5580(.077)[.086]	.5084(.089)[.103]	.5615(.050)	.5483(.049)[.057]	.5017(.056)[.065]
λ	.1560(.046)	.1641(.046)[.048]	.1944(.054)[.062]	.1620(.029)	.1701(.029)[.032]	.1990(.033)[.039]
σ_v^2	.7777(.057)	.9993(.073)[.076]	–	.8905(.047)	1.0041(.053)[.053]	–

Table 5: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP3` with **homoskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \lambda_1, \lambda_2, \rho_1, \rho_2, \sigma_v^2) = (1, 0.2, 0.3, 0.2, 0.3, 1)$, $\mathtt{W}_1 = \mathtt{M}_1 = \mathtt{Queen}$, $\mathtt{W}_2 = \mathtt{M}_2 = \mathtt{Rook}$.

	T=5			**T=10**		
	QMLE	M-Est	M-Est	QMLE	M-Est	RM-Est
	$n = 50$; error = 1, 2, 3, for the three panels below					
β_1	.9990(.037)	.9968(.037)[.038]	.9969(.037)[.037]	.9989(.023)	.9986(.023)[.024]	.9986(.023)[.024]
λ_1	.2018(.079)	.1855(.085)[.081]	.1858(.085)[.080]	.1983(.046)	.1962(.050)[.048]	.1964(.050)[.048]
λ_2	.2985(.061)	.2932(.061)[.061]	.2933(.061)[.061]	.2981(.036)	.2988(.037)[.038]	.2988(.037)[.038]
ρ_1	.0492(.253)	.2017(.211)[.185]	.2018(.212)[.182]	.1003(.117)	.1983(.115)[.108]	.1982(.116)[.108]
ρ_2	.3037(.178)	.3122(.153)[.141]	.3126(.153)[.140]	.2714(.091)	.2979(.087)[.083]	.2980(.087)[.083]
σ_v^2	.7134(.082)	.9489(.108)[.106]	–	.8601(.060)	.9820(.069)[.071]	–
β_1	.9972(.038)	.9949(.038)[.038]	.9951(.038)[.037]	.9989(.024)	.9983(.024)[.024]	.9983(.024)[.023]
λ_1	.2012(.076)	.1861(.083)[.081]	.1864(.083)[.080]	.1998(.048)	.1955(.052)[.048]	.1958(.051)[.047]
λ_2	.2994(.060)	.2943(.059)[.060]	.2949(.059)[.059]	.2975(.037)	.2971(.038)[.038]	.2970(.038)[.037]
ρ_1	.0657(.239)	.2118(.208)[.184]	.2122(.209)[.180]	.0928(.122)	.1996(.127)[.109]	.1993(.125)[.106]
ρ_2	.3045(.170)	.3144(.148)[.138]	.3135(.147)[.136]	.2775(.088)	.3064(.090)[.083]	.3069(.090)[.081]
σ_v^2	.7210(.177)	.9575(.236)[.210]	–	.8579(.139)	.9800(.159)[.154]	–
β_1	.9970(.038)	.9952(.038)[.038]	.9952(.038)[.037]	.9985(.023)	.9981(.023)[.024]	.9981(.023)[.024]
λ_1	.2007(.077)	.1853(.084)[.082]	.1854(.084)[.080]	.1992(.044)	.1962(.049)[.049]	.1960(.049)[.048]
λ_2	.2980(.063)	.2931(.060)[.060]	.2932(.060)[.060]	.2981(.037)	.2980(.039)[.038]	.2981(.039)[.038]
ρ_1	.0455(.246)	.1978(.212)[.185]	.1994(.212)[.182]	.0948(.116)	.1961(.119)[.110]	.1965(.119)[.108]
ρ_2	.3019(.176)	.3099(.150)[.140]	.3107(.149)[.138]	.2765(.087)	.3041(.087)[.083]	.3039(.087)[.082]
σ_v^2	.7221(.134)	.9609(.178)[.159]	–	.8639(.105)	.9862(.121)[.115]	–
	$n = 100$; error = 1, 2, 3, for the three panels below					
β_1	.9979(.027)	.9977(.027)[.028]	.9977(.027)[.028]	.9995(.018)	.9995(.018)[.018]	.9995(.018)[.018]
λ_1	.1998(.053)	.1953(.055)[.052]	.1954(.055)[.052]	.1981(.036)	.1967(.037)[.037]	.1966(.037)[.036]
λ_2	.2970(.044)	.2982(.043)[.042]	.2983(.044)[.042]	.2996(.029)	.3002(.030)[.029]	.3003(.030)[.029]
ρ_1	.1573(.156)	.1988(.131)[.122]	.1987(.131)[.121]	.1568(.081)	.1969(.076)[.076]	.1972(.076)[.075]
ρ_2	.3417(.119)	.2968(.100)[.096]	.2972(.100)[.095]	.3034(.064)	.2997(.059)[.059]	.2998(.059)[.058]
σ_v^2	.7451(.062)	.9794(.080)[.077]	–	.8734(.045)	.9907(.051)[.050]	–
β_1	.9992(.029)	.9990(.029)[.028]	.9990(.029)[.028]	.9980(.018)	.9980(.018)[.018]	.9980(.018)[.017]
λ_1	.1989(.052)	.1946(.054)[.052]	.1943(.054)[.051]	.2006(.035)	.1994(.036)[.036]	.1994(.036)[.036]
λ_2	.2976(.041)	.2990(.040)[.042]	.2992(.040)[.041]	.2990(.028)	.2999(.029)[.029]	.2999(.029)[.029]
ρ_1	.1438(.157)	.1877(.129)[.123]	.1881(.130)[.122]	.1535(.082)	.1937(.076)[.076]	.1938(.076)[.075]
ρ_2	.3457(.114)	.2988(.096)[.095]	.2989(.096)[.093]	.3011(.065)	.2973(.060)[.059]	.2976(.060)[.058]
σ_v^2	.7443(.124)	.9787(.163)[.158]	–	.8723(.102)	.9895(.116)[.112]	–
β_1	.9990(.027)	.9988(.027)[.028]	.9988(.027)[.027]	1.0003(.018)	1.0003(.018)[.018]	1.0003(.018)[.017]
λ_1	.2033(.052)	.1987(.055)[.051]	.1988(.055)[.051]	.2007(.035)	.1994(.036)[.036]	.1995(.036)[.036]
λ_2	.2970(.043)	.2985(.042)[.043]	.2986(.042)[.042]	.2982(.027)	.2990(.028)[.029]	.2989(.028)[.029]
ρ_1	.1454(.150)	.1892(.127)[.122]	.1896(.127)[.121]	.1554(.081)	.1952(.075)[.076]	.1951(.075)[.075]
ρ_2	.3504(.115)	.3031(.097)[.096]	.3033(.097)[.094]	.3002(.061)	.2966(.057)[.059]	.2968(.057)[.059]
σ_v^2	.7462(.094)	.9811(.123)[.116]	–	.8720(.070)	.9892(.080)[.081]	–

Table 5: Cont'd

	T=5			T=10		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
	$n = 200$; error = 1, 2, 3, for the three panels below					
β_1	.9991(.020)	.9991(.020)[.019]	.9991(.020)[.019]	.9994(.013)	.9994(.013)[.012]	.9994(.013)[.012]
λ_1	.2013(.034)	.1993(.035)[.034]	.1993(.035)[.033]	.1995(.025)	.1986(.026)[.026]	.1986(.026)[.025]
λ_2	.2971(.026)	.2992(.026)[.027]	.2992(.026)[.027]	.2993(.018)	.2999(.019)[.019]	.2999(.019)[.019]
ρ_1	.1980(.103)	.1969(.093)[.083]	.1970(.094)[.083]	.1907(.056)	.2014(.052)[.053]	.2014(.052)[.053]
ρ_2	.3660(.075)	.2988(.068)[.065]	.2991(.068)[.065]	.3163(.045)	.2974(.042)[.040]	.2975(.042)[.041]
σ_v^2	.7528(.044)	.9902(.057)[.055]	–	.8763(.033)	.9942(.037)[.036]	–
β_1	.9998(.019)	.9997(.019)[.019]	.9997(.019)[.019]	.9996(.012)	.9996(.012)[.012]	.9996(.012)[.012]
λ_1	.1999(.033)	.1978(.035)[.033]	.1977(.035)[.033]	.2002(.025)	.1993(.025)[.025]	.1993(.025)[.025]
λ_2	.2984(.026)	.3003(.026)[.027]	.3003(.026)[.027]	.2987(.019)	.2992(.019)[.019]	.2993(.019)[.019]
ρ_1	.2026(.098)	.2013(.091)[.082]	.2016(.091)[.081]	.1867(.056)	.1978(.052)[.053]	.1980(.052)[.053]
ρ_2	.3694(.077)	.3025(.070)[.065]	.3027(.070)[.064]	.3186(.045)	.2994(.041)[.041]	.2994(.041)[.040]
σ_v^2	.7504(.091)	.9873(.119)[.114]	–	.8766(.069)	.9946(.078)[.080]	–
β_1	.9991(.019)	.9990(.019)[.019]	.9991(.019)[.019]	.9997(.013)	.9997(.013)[.012]	.9997(.013)[.012]
λ_1	.1992(.033)	.1965(.037)[.034]	.1965(.037)[.033]	.2017(.026)	.2009(.027)[.025]	.2009(.027)[.025]
λ_2	.2983(.027)	.2998(.027)[.027]	.2997(.027)[.027]	.2992(.018)	.2998(.018)[.019]	.2999(.018)[.019]
ρ_1	.2008(.105)	.2018(.099)[.082]	.2018(.099)[.082]	.1870(.060)	.1980(.056)[.054]	.1981(.056)[.053]
ρ_2	.3637(.074)	.2990(.070)[.065]	.2990(.070)[.065]	.3169(.043)	.2978(.040)[.041]	.2978(.040)[.040]
σ_v^2	.7530(.065)	.9902(.086)[.084]	–	.8771(.052)	.9951(.059)[.058]	–
	$n = 400$; error = 1, 2, 3, for the three panels below					
β_1	.9995(.014)	.9998(.014)[.014]	.9998(.014)[.013]	.9998(.009)	.9999(.009)[.009]	.9999(.009)[.009]
λ_1	.1999(.025)	.2000(.025)[.024]	.2000(.025)[.024]	.1987(.018)	.2001(.018)[.018]	.1999(.018)[.018]
λ_2	.2964(.019)	.2997(.019)[.019]	.2996(.019)[.019]	.2974(.013)	.2999(.014)[.014]	.3000(.014)[.014]
ρ_1	.2236(.071)	.1972(.061)[.058]	.1973(.061)[.057]	.2058(.041)	.1996(.038)[.037]	.1997(.038)[.037]
ρ_2	.3791(.053)	.2994(.046)[.045]	.2995(.046)[.045]	.3446(.031)	.3001(.028)[.028]	.3001(.028)[.028]
σ_v^2	.7577(.031)	.9946(.040)[.039]	–	.8735(.023)	.9983(.025)[.025]	–
β_1	.9995(.014)	.9997(.014)[.014]	.9997(.014)[.013]	1.0001(.009)	1.0002(.009)[.009]	1.0002(.009)[.009]
λ_1	.1982(.025)	.1983(.025)[.024]	.1983(.025)[.024]	.2001(.018)	.1997(.018)[.019]	.1997(.018)[.018]
λ_2	.2965(.018)	.2998(.018)[.019]	.2999(.018)[.019]	.2985(.013)	.2991(.014)[.014]	.2991(.014)[.014]
ρ_1	.2273(.070)	.2003(.060)[.058]	.2004(.059)[.057]	.2129(.039)	.2010(.038)[.038]	.2010(.038)[.037]
ρ_2	.3792(.052)	.2996(.045)[.045]	.2999(.045)[.045]	.3332(.033)	.3015(.029)[.029]	.3016(.028)[.029]
σ_v^2	.7583(.063)	.9955(.082)[.082]	–	.9054(.052)	1.0001(.057)[.059]	–
β_1	.9999(.014)	1.0002(.014)[.014]	1.0002(.014)[.013]	.9994(.009)	.9994(.009)[.009]	.9994(.009)[.009]
λ_1	.2003(.024)	.2007(.024)[.024]	.2007(.024)[.024]	.2001(.017)	.1996(.018)[.018]	.1996(.018)[.018]
λ_2	.2967(.019)	.3001(.019)[.019]	.3002(.019)[.019]	.2987(.013)	.2995(.014)[.014]	.2993(.014)[.014]
ρ_1	.2221(.068)	.1954(.057)[.058]	.1955(.057)[.058]	.1966(.041)	.1998(.038)[.037]	.1995(.037)[.036]
ρ_2	.3774(.054)	.2978(.046)[.045]	.2977(.046)[.045]	.3345(.029)	.3001(.028)[.029]	.3004(.028)[.029]
σ_v^2	.7580(.047)	.9949(.061)[.060]	–	.8696(.036)	.9985(.041)[.040]	–

Table 6: Empirical mean(sd)[$\hat{se}$] of QMLE, M-estimator and RM-estimator: `DGP3` with **heteroskedasticity**, Unbalancedness percentage = 10%, $(\beta_1, \lambda_1, \lambda_2, \rho_1, \rho_2, \sigma_v^2) = (1, 0.2, 0.3, 0.2, 0.3, 1)$, $\mathtt{W}_1 = \mathtt{M}_1 =$ `Group-II`, $\mathtt{W}_2 = \mathtt{M}_2 =$ `Rook`.

	T=5			**T=10**		
	QMLE	M-Est	M-Est	QMLE	M-Est	RM-Est
	$n = 50$; error = 1, 2, 3, for the three panels below					
β_1	.9976(.043)	.9983(.043)[.041]	.9984(.044)[.043]	1.0011(.025)	1.0006(.025)[.026]	1.0004(.025)[.025]
λ_1	.1758(.065)	.1855(.065)[.080]	.1952(.082)[.085]	.1789(.033)	.1885(.032)[.045]	.1998(.037)[.038]
λ_2	.2953(.054)	.2977(.055)[.064]	.2977(.055)[.058]	.2964(.038)	.2993(.039)[.038]	.2987(.039)[.038]
ρ_1	-.0559(.180)	.0933(.126)[.174]	.1236(.276)[.301]	-.0199(.103)	.0963(.087)[.116]	.1584(.175)[.173]
ρ_2	.3181(.175)	.2991(.143)[.143]	.2935(.140)[.137]	.2851(.090)	.2999(.082)[.084]	.3008(.082)[.083]
σ_v^2	.7162(.099)	.9576(.130)[.137]	–	.8419(.074)	.9665(.085)[.091]	–
β_1	.9975(.043)	.9979(.043)[.041]	.9982(.044)[.042]	1.0006(.026)	1.0000(.026)[.026]	.9999(.026)[.025]
λ_1	.1724(.068)	.1806(.068)[.081]	.1890(.086)[.086]	.1785(.031)	.1878(.031)[.045]	.1998(.036)[.037]
λ_2	.2932(.059)	.2942(.060)[.064]	.2948(.059)[.058]	.2963(.037)	.2991(.038)[.038]	.2987(.038)[.038]
ρ_1	-.0532(.192)	.1052(.131)[.175]	.1498(.271)[.289]	-.0263(.103)	.0923(.086)[.119]	.1527(.171)[.172]
ρ_2	.3222(.173)	.3065(.145)[.143]	.3000(.138)[.133]	.2824(.089)	.2974(.083)[.084]	.2982(.082)[.081]
σ_v^2	.7260(.221)	.9706(.295)[.264]	–	.8433(.175)	.9682(.201)[.189]	–
β_1	.9995(.044)	.9999(.043)[.041]	1.0004(.044)[.043]	.9997(.024)	.9991(.024)[.026]	.9988(.025)[.025]
λ_1	.1734(.068)	.1813(.067)[.082]	.1925(.081)[.086]	.1795(.033)	.1885(.032)[.045]	.1988(.039)[.039]
λ_2	.2923(.058)	.2931(.060)[.064]	.2936(.059)[.058]	.2956(.039)	.2982(.040)[.038]	.2978(.040)[.038]
ρ_1	-.0582(.176)	.0990(.137)[.174]	.1315(.277)[.299]	-.0179(.098)	.0994(.086)[.117]	.1653(.172)[.174]
ρ_2	.3168(.174)	.3035(.146)[.144]	.2968(.139)[.135]	.2849(.090)	.3008(.084)[.083]	.3012(.083)[.082]
σ_v^2	.7227(.155)	.9665(.207)[.197]	–	.8414(.123)	.9659(.141)[.139]	–
	$n = 100$; error = 1, 2, 3, for the three panels below					
β_1	.9980(.029)	.9972(.029)[.028]	.9983(.030)[.029]	.9992(.018)	1.0000(.018)[.018]	.9999(.018)[.018]
λ_1	.1922(.044)	.1856(.046)[.061]	.1963(.054)[.052]	.1942(.027)	.1950(.022)[.032]	.2003(.025)[.025]
λ_2	.2966(.038)	.2952(.041)[.039]	.2964(.039)[.040]	.2974(.028)	.2995(.028)[.028]	.2993(.028)[.029]
ρ_1	.0382(.191)	.1294(.130)[.111]	.1754(.193)[.178]	.0534(.177)	.1284(.071)[.076]	.1843(.109)[.104]
ρ_2	.3479(.109)	.3069(.109)[.091]	.3036(.100)[.092]	.3080(.066)	.2983(.057)[.059]	.2977(.057)[.059]
σ_v^2	.7486(.076)	.9790(.100)[.102]	–	.8728(.055)	.9914(.061)[.067]	–
β_1	.9986(.028)	.9976(.029)[.028]	.9987(.030)[.029]	.9991(.018)	.9996(.018)[.018]	.9995(.018)[.018]
λ_1	.1893(.042)	.1832(.047)[.061]	.1945(.054)[.053]	.1926(.025)	.1940(.022)[.032]	.1994(.025)[.025]
λ_2	.2964(.038)	.2945(.041)[.038]	.2955(.040)[.039]	.2977(.029)	.2997(.029)[.028]	.2995(.029)[.029]
ρ_1	.0434(.179)	.1324(.135)[.114]	.1797(.202)[.177]	.0577(.148)	.1251(.070)[.077]	.1801(.107)[.103]
ρ_2	.3518(.105)	.3131(.110)[.092]	.3088(.099)[.090]	.3121(.064)	.3020(.059)[.059]	.3015(.059)[.057]
σ_v^2	.7517(.161)	.9832(.211)[.205]	–	.8678(.128)	.9865(.145)[.142]	–
β_1	.9986(.028)	.9973(.029)[.028]	.9984(.029)[.029]	.9986(.018)	.9992(.017)[.018]	.9991(.017)[.018]
λ_1	.1895(.043)	.1822(.047)[.060]	.1933(.054)[.052]	.1926(.025)	.1943(.022)[.032]	.1998(.024)[.024]
λ_2	.2945(.039)	.2923(.042)[.038]	.2934(.041)[.039]	.2964(.028)	.2983(.028)[.028]	.2981(.028)[.029]
ρ_1	.0390(.183)	.1313(.135)[.112]	.1770(.197)[.176]	.0570(.146)	.1233(.069)[.077]	.1774(.106)[.104]
ρ_2	.3531(.109)	.3144(.112)[.091]	.3105(.101)[.090]	.3087(.064)	.2991(.057)[.059]	.2987(.057)[.058]
σ_v^2	.7448(.115)	.9742(.150)[.150]	–	.8696(.093)	.9885(.105)[.103]	–

Table 6: Cont'd.

	T=5			**T=10**		
	QMLE	M-Est	RM-Est	QMLE	M-Est	RM-Est
	$n = 200$; error = 1, 2, 3, for the three panels below					
β_1	1.0009(.019)	1.0012(.019)[.020]	1.0011(.019)[.019]	1.0002(.012)	1.0003(.012)[.012]	1.0001(.012)[.012]
λ_1	.1982(.029)	.1970(.027)[.033]	.2013(.030)[.031]	.1926(.016)	.1937(.016)[.022]	.1993(.018)[.019]
λ_2	.2965(.028)	.3004(.028)[.028]	.3001(.028)[.028]	.2987(.019)	.3001(.019)[.020]	.3000(.019)[.019]
ρ_1	.1005(.144)	.1290(.069)[.074]	.1809(.103)[.104]	.1014(.072)	.1268(.050)[.054]	.1885(.076)[.075]
ρ_2	.3721(.076)	.2970(.064)[.065]	.2981(.065)[.064]	.3236(.044)	.3010(.041)[.041]	.2999(.041)[.041]
σ_v^2	.7581(.048)	.9924(.061)[.066]	–	.8739(.039)	.9914(.044)[.048]	–
β_1	.9977(.020)	.9982(.019)[.020]	.9981(.019)[.019]	.9997(.013)	.9998(.012)[.012]	.9996(.012)[.012]
λ_1	.1976(.031)	.1964(.028)[.033]	.2007(.031)[.031]	.1931(.017)	.1942(.017)[.022]	.1997(.019)[.019]
λ_2	.2952(.028)	.2994(.028)[.028]	.2990(.029)[.028]	.2994(.018)	.3007(.018)[.020]	.3007(.018)[.019]
ρ_1	.0956(.143)	.1252(.069)[.076]	.1756(.102)[.103]	.1036(.063)	.1278(.049)[.055]	.1900(.074)[.074]
ρ_2	.3733(.073)	.2976(.061)[.065]	.2993(.062)[.064]	.3218(.044)	.2994(.041)[.041]	.2985(.041)[.040]
σ_v^2	.7607(.105)	.9956(.137)[.137]	–	.8711(.093)	.9882(.105)[.103]	–
β_1	.9991(.019)	.9996(.019)[.020]	.9994(.019)[.019]	1.0003(.012)	1.0003(.012)[.012]	1.0001(.012)[.012]
λ_1	.1962(.031)	.1949(.029)[.033]	.1987(.032)[.031]	.1933(.016)	.1945(.016)[.022]	.1999(.018)[.019]
λ_2	.2978(.029)	.3015(.029)[.028]	.3012(.029)[.028]	.2992(.020)	.3004(.020)[.020]	.3003(.020)[.019]
ρ_1	.0992(.139)	.1284(.074)[.075]	.1795(.108)[.103]	.1067(.054)	.1297(.050)[.054]	.1931(.076)[.074]
ρ_2	.3707(.077)	.2972(.067)[.065]	.2982(.066)[.064]	.3212(.045)	.2991(.041)[.041]	.2982(.041)[.040]
σ_v^2	.7539(.075)	.9866(.098)[.100]	–	.8728(.064)	.9901(.073)[.075]	–
	$n = 400$; error = 1, 2, 3, for the three panels below					
β_1	.9979(.013)	.9986(.013)[.013]	.9988(.013)[.013]	1.0017(.009)	1.0017(.009)[.009]	.9999(.009)[.009]
λ_1	.1936(.021)	.1924(.018)[.023]	.2004(.020)[.020]	.1913(.012)	.1919(.016)[.016]	.2002(.013)[.014]
λ_2	.2932(.021)	.2981(.020)[.020]	.2987(.020)[.020]	.2990(.015)	.3002(.015)[.014]	.3006(.015)[.014]
ρ_1	.0852(.146)	.1030(.053)[.055]	.1873(.098)[.094]	.1310(.035)	.1364(.037)[.036]	.1976(.050)[.049]
ρ_2	.3966(.056)	.3051(.049)[.047]	.3020(.048)[.047]	.3134(.033)	.2853(.030)[.029]	.3008(.029)[.029]
σ_v^2	.7592(.036)	1.0021(.047)[.052]	–	.8776(.026)	.9944(.033)[.034]	–
β_1	.9986(.014)	.9991(.014)[.013]	.9994(.014)[.013]	1.0019(.009)	1.0020(.009)[.009]	.9994(.009)[.009]
λ_1	.1924(.018)	.1917(.017)[.023]	.1996(.020)[.020]	.1977(.013)	.1977(.016)[.016]	.1996(.013)[.014]
λ_2	.2946(.020)	.2994(.020)[.020]	.3000(.020)[.020]	.2920(.016)	.2929(.015)[.014]	.2995(.015)[.014]
ρ_1	.0950(.118)	.1058(.051)[.055]	.1925(.091)[.092]	.1392(.037)	.1451(.037)[.037]	.1987(.050)[.049]
ρ_2	.3916(.055)	.3016(.049)[.047]	.2983(.047)[.046]	.3375(.034)	.3078(.030)[.029]	.2990(.029)[.029]
σ_v^2	.7599(.082)	1.0027(.107)[.108]	–	.8666(.064)	.9827(.072)[.072]	–
β_1	.9992(.014)	.9998(.013)[.013]	1.0001(.013)[.013]	.9993(.009)	.9994(.009)[.009]	.9994(.009)[.009]
λ_1	.1921(.020)	.1912(.018)[.023]	.1995(.020)[.020]	.1955(.014)	.1958(.016)[.016]	.1997(.013)[.014]
λ_2	.2937(.019)	.2986(.019)[.020]	.2990(.019)[.020]	.2924(.015)	.2937(.015)[.014]	.2999(.015)[.014]
ρ_1	.0872(.140)	.1028(.047)[.055]	.1882(.088)[.093]	.1227(.038)	.1292(.037)[.037]	.1964(.050)[.050]
ρ_2	.3917(.056)	.3007(.047)[.047]	.2983(.047)[.046]	.3392(.031)	.3086(.030)[.029]	.3007(.029)[.029]
σ_v^2	.7628(.059)	1.0066(.077)[.081]	–	.8581(.044)	.9732(.052)[.049]	–

References

Andrews, D. W. (1992). Generic uniform convergence. *Econometric theory 8*, 241–257.

Anselin, L., J. L. Gallo, and H. Jayet (2008). Spatial Panel Econometrics. In L. Mátyás and P. Sevestre (Eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Advanced Studies in Theoretical and Applied Econometrics, pp. 625–660. Berlin, Heidelberg: Springer.

Davidson, J. (1994). *Stochastic limit theory: An introduction for econometricians*. OUP Oxford.

Feller, W. (1967). An introduction to probability theory and its applications. *1957*.

Kelejian, H. H. and I. R. Prucha (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review 40*, 509–533.

Kelejian, H. H. and I. R. Prucha (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics 104*, 219–257.

Lee, L.-F. (2002). Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. *Econometric theory 18*, 252–277.

Lee, L.-F. (2004). A supplement to 'asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models'. *https://www.asc.ohio-state.edu/lee.1777/wp/sar-qml-r-appen-04feb.pdf*.

Lee, L.-F. and J. Yu (2016). Identification of spatial Durbin panel models. *Journal of Applied Econometrics 31*, 133–162.

Lin, X. and L.-F. Lee (2010). GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics 157*, 34–52.

Liu, S. F. and Z. Yang (2015). Modified QML estimation of spatial autoregressive models with unknown heteroskedasticity and nonnormality. *Regional Science and Urban Economics 52*, 50–70.

Liu, S. F. and Z. Yang (2020). Robust estimation and inference of spatial panel data models with fixed effects. *Japanese Journal of Statistics and Data Science 3*, 257–311.

Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Varah, J. M. (1975). A lower bound for the smallest singular value of a matrix. *Linear Algebra and its applications 11*, 3–5.

Yang, Z. (2015). A general method for third-order bias and variance corrections on a nonlinear estimator. *Journal of Econometrics 186*, 178–200.