Online Supplementary Appendices for "Simultaneous Equations Models with Higher-Order Spatial or Social Network Interactions"[a]

David M. Drukker,[b] Peter H. Egger,[c] and Ingmar R. Prucha[d]

December 27, 2021 (This Version)

[a]We are grateful to the editor, a guest editor, and three anonymous referees for their very helpful comments. We gratefully acknowledge financial support from the National Institute of Health through SBIR grants R43 AG027622 and R44 AG027622. We also thank the the CESifo for their hospitality.

[b]Department of Economics and International Business, Sam Houston State University, Huntsville, TX 77351, Tel.:+1-936-294-4791, e-mail: dmetrics@gmail.com

[c]KOF, ETH Zurich, Leonhardstrasse 21, 8092 Zurich, Switzerland; Tel.: +41-44-632-4108, e-mail: egger@kof.ethz.ch

[d]Department of Economics, University of Maryland, College Park, MD 20742, Tel.: +1-301-405-3499, e-mail: prucha@umd.edu

Abstract

This supplement provides additional technical material, expanded proofs, and reports on a Monte Carlo study that explores the small sample properties of our estimators. For the main Assumptions, notations and definitions of various quantities appearing in the sequel, the reader is referred to the main paper.

D Supplement to Appendix A

Proof of Lemma A.1: Let $C_A = \sup_n \max_{i=1}^{m_n} \sum_{j=1}^{n} |a_{ij,n}|$, $C_\mu = \sup_n \max_{i=1}^{m_n} E|\mu_{i,n}|^p$ and $C_\eta = \sup_n \max_{i=1}^{m_n} E|\eta_{i,n}|^p$. Clearly

$$|\xi_{i,n}| \leq |\mu_{i,n}| + \sum_{j=1}^{m_n} |a_{ij,n}| |\eta_{jn}| \leq |\mu_{i,n}| + C_A \sum_{j=1}^{m_n} b_{ij,n} |\eta_{jn}|$$

with $b_{ij,n} = |a_{ij,n}| / \left(\sum_{j=1}^{m_n} |a_{ij,n}|\right)$ if $\sum_{j=1}^{m_n} |a_{ij,n}| > 0$ and $b_{ij,n} = 0$ if $\sum_{j=1}^{m_n} |a_{ij,n}| = 0$. Clearly $E|\xi_{i,n}|^p \leq C_\mu$ for $\sum_{j=1}^{m_n} |a_{ij,n}| = 0$. For $\sum_{j=1}^{m_n} |a_{ij,n}| > 0$ observe that since $0 \leq b_{ij,n} \leq 1$ and $\sum_{j=1}^{m_n} b_{ij,n} = 1$ it follows from Holder's inequality that $\sum_{j=1}^{m_n} b_{ij,n} |\eta_{j,n}| \leq \left[\sum_{j=1}^{m_n} b_{ij,n} |\eta_{j,n}|^p\right]^{1/p}$ and consequently

$$\begin{aligned} E|\xi_{i,n}|^p &\leq 2^p E|\mu_{i,n}|^p + 2^p C_A^p E\left[\sum_{j=1}^{m_n} b_{ij,n} |\eta_{j,n}|\right]^p \\ &\leq 2^p E|\mu_{i,n}|^p + 2^p C_A^p \sum_{j=1}^{m_n} b_{ij,n} E[|\eta_{j,n}|^p] \leq 2^p C_\mu + 2^p C_A^p C_\eta < \infty \end{aligned}$$

which proves the claim since C_A, C_μ and C_η do not depend on i and n. ■

Proof of Lemma A.2: By assumption $\mathbf{X}_n$ is non-stochastic with $\sup_n \sup_{i,k} |x_{ik,n}| < \infty$, and so (A.1) holds trivially if $z_{ij,n}$ corresponds to an element of $\mathbf{X}_n$. Next observe that by (7) and (9) we have

$$\begin{aligned} \mathbf{y}_n &= \mathbf{a}_n + \mathbf{A}_n \boldsymbol{\nu}_n, \\ \mathbf{a}_n &= (\mathbf{I}_{nG} - \mathbf{B}_n^*)^{-1} \mathbf{C}_n^* \mathbf{x}_n, \\ \mathbf{A}_n &= (\mathbf{I}_{nG} - \mathbf{B}_n^*)^{-1} (\mathbf{I}_{nG} - \mathbf{R}_n^*)^{-1} (\boldsymbol{\Sigma}_\star' \otimes \mathbf{I}_n). \end{aligned}$$

In light of Assumptions 1-3 the absolute elements of $\mathbf{a}_n$ are uniformly bounded, and furthermore the row and column sums of the absolute elements of $\mathbf{A}_n$ are uniformly bounded; compare, e.g., Remark A.1 in Kelejian and Prucha (2004). By Assumption 4 the elements of $\boldsymbol{\nu}_n$ are i.i.d. with finite fourth moments. Thus it follows immediately from Lemma A.1 that $\sup_n \sup_{i,l} E|y_{il,n}|^4 < \infty$. Next observe that the columns of $\overline{\mathbf{Y}}_n$ are of the form $\overline{\mathbf{y}}_{l,s,n} = \mathbf{W}_{s,n} \mathbf{y}_{l,n}$. Since by Assumption 1 the row and column sums of the absolute elements of $\mathbf{W}_{s,n}$ are uniformly bounded it follows further from Lemma A.1 that $\sup_n \sup_{i,l,s} E\left|\overline{y}_{il,s,n}\right|^4 < \infty$, which completes the proof. ■

Proof of Lemma A.3: In light of the proof of Lemma A.2, and observing that $\mathbf{u}_n = (\mathbf{I}_{nG} - \mathbf{R}_n^*)^{-1}(\mathbf{\Sigma}_*' \otimes \mathbf{I}_n)\boldsymbol{\nu}_n$, it is readily seen that under the maintained assumptions $\mathbf{u}_{g,n}$ and all columns of $\underline{\mathbf{Z}}_n$ are of the generic form

$$\mathbf{C}_{g,n}\boldsymbol{\nu}_n,\ \mathbf{c}_{g,n} \text{ or } \mathbf{c}_{g,n} + \mathbf{C}_{g,n}\boldsymbol{\nu}_n, \tag{D.1}$$

where $\mathbf{c}_{g,n}$ is an $n \times 1$ nonstochastic vector with uniformly bounded elements and $\mathbf{C}_{g,n}$ is an $n \times nG$ nonstochastic matrix whose row and column sums are uniformly bounded in absolute value. By Assumption 4 the elements of the $nG \times 1$ vector $\boldsymbol{\nu}_n$ are i.i.d. $(0,1)$ with finite fourth moments. Given this, it is readily seen that $n^{-1}\mathbf{u}_{h,n}'\mathbf{A}_n\mathbf{u}_{g,n}$ and the elements of $n^{-1}\underline{\mathbf{Z}}_n'\mathbf{A}_n\mathbf{u}_{g,n}$ and $n^{-1}\underline{\mathbf{Z}}_n\mathbf{A}_n\underline{\mathbf{Z}}_n$ are of the generic form d_n, $n^{-1}\mathbf{d}_n'\boldsymbol{\nu}_n$ or $n^{-1}\nu_n'\mathbf{D}_n\boldsymbol{\nu}_n$, or sums thereof, where $|d_n|$, the absolute elements of $\mathbf{d}_n$ and the row and column sums of the absolute elements of $\mathbf{D}_n$ are uniformly bounded by some finite constant, say K. In the following let $\overline{\mathbf{D}}_n = (\overline{d}_{lk,n}) = (\mathbf{D}_n + \mathbf{D}_n')/2$. Observe that $E\left[n^{-1}\mathbf{d}_n'\boldsymbol{\nu}_n\right] = 0$ and $E\left[n^{-1}\boldsymbol{\nu}_n'\mathbf{D}_n\boldsymbol{\nu}_n\right] = n^{-1}tr(\mathbf{D}_n) = O(1)$. Furthermore observe that $var(n^{-1}\mathbf{d}_n'\boldsymbol{\nu}_n) \leq n^{-1}K^2 = o(1)$, and that in light of, e.g., Lemma A.1 in Kelejian and Prucha (2004) and Remark A.2 in Kapoor et al. (2007), we have $var(n^{-1}\boldsymbol{\nu}_n'\mathbf{D}_n\boldsymbol{\nu}_n) \leq n^{-2}tr(\overline{\mathbf{D}}_n^2) + n^{-2}K\sum_{i,g}\left|Ev_{ig,n}^4 - 3\right| \leq n^{-1}K_* = o(1)$ for some finite constant K_*. Thus clearly $n^{-1}\mathbf{u}_{h,n}'\mathbf{A}_n\mathbf{u}_{g,n} = O_p(1)$, $n^{-1}\underline{\mathbf{Z}}_n\mathbf{A}_n\mathbf{u}_{g,n} = O_p(1)$ and $n^{-1}\underline{\mathbf{Z}}_n\mathbf{A}_n\underline{\mathbf{Z}}_n = O_p1)$. The third claim in the lemma follows from Chebychev's inequality. ■

Proof of Lemma A.4: Clearly,

$$n^{1/2}(\widetilde{\boldsymbol{\delta}}_{g,n} - \boldsymbol{\delta}_{g,n}) = \widetilde{\mathbf{P}}_{gg,n}' n^{-1/2}\mathbf{F}_{gg,n}'\boldsymbol{\varepsilon}_{g,n},$$

where $\widetilde{\mathbf{P}}_{gg,n}$ and $\mathbf{F}_{gg,n}$ are defined in the lemma. Given Assumption 6 clearly $\widetilde{\mathbf{P}}_{gg,n} = \mathbf{P}_{gg} + o_p(1)$ with $\mathbf{P}_{gg}$ being finite, which establishes (c). Since by Assumption 2 the row and column sums of $(\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime})^{-1}$ are uniformly bounded in absolute value, and since by Assumption 5 the elements of $\mathbf{H}_n$ are uniformly bounded in absolute value, it follows that the elements of $\mathbf{F}_{gg,n}$ are uniformly bounded in absolute value. By Assumption 4, $E(\boldsymbol{\varepsilon}_{g,n}) = 0$ and $E(\boldsymbol{\varepsilon}_{g,n}\boldsymbol{\varepsilon}_{g,n}') = \sigma_{gg}\mathbf{I}_n$. Therefore, $En^{-1/2}\mathbf{F}_{gg,n}'\boldsymbol{\varepsilon}_{g,n} = \mathbf{0}$ and the elements of

$VC(n^{-1/2}\mathbf{F}_n'\boldsymbol{\varepsilon}_{g,n}) = \sigma_{gg}n^{-1}\mathbf{F}_{gg,n}'\mathbf{F}_{gg,n}$ are also uniformly bounded in absolute value. Thus, by Chebyshev's inequality $n^{-1/2}\mathbf{F}_{gg,n}'\boldsymbol{\varepsilon}_{g,n} = O_p(1)$, and consequently $n^{1/2}(\widetilde{\boldsymbol{\delta}}_{g,n} - \boldsymbol{\delta}_{g,n}) = \mathbf{P}_{gg}'n^{-1/2}\mathbf{F}_{gg,n}'\boldsymbol{\varepsilon}_{g,n} + o_p(1)$ and $\mathbf{P}_{gg}'n^{-1/2}\mathbf{F}_{gg,n}'\boldsymbol{\varepsilon}_{g,n} = O_p(1)$. This establishes (a) and (b), recalling that $\mathbf{T}_{gg,n} = \mathbf{F}_{gg,n}\mathbf{P}_{gg,n}$. Next observe that

$$\begin{aligned}
& \lambda_{\min}(n^{-1}\mathbf{T}_{gg,n}'\mathbf{T}_{gg,n}) \\
\geq \; & \lambda_{\min}\left[\left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime}\right)^{-1}\left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*}\right)^{-1}\right]\lambda_{\min}\left[n^{-1}\mathbf{H}_n'\mathbf{H}_n\right] \\
& \lambda_{\min}\left\{[\mathbf{Q}_{\mathbf{HZ},g}'\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g}]^{-1}\mathbf{Q}_{\mathbf{HZ},g}'\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g}[\mathbf{Q}_{\mathbf{HZ},g}'\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g}]^{-1}\right\} \\
\geq \; & c
\end{aligned}$$

for some $c > 0$, since in light of Assumptions 1 and 2 the largest eigenvalue of $\left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*}\right)\left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime}\right)$ is bounded from above, and thus the smallest eigenvalue of $\left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime}\right)^{-1}\left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*}\right)^{-1}$ is bounded away from zero, and since $\lambda_{\min}\left[n^{-1}\mathbf{H}_n'\mathbf{H}_n\right] \geq \left[\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}})\right]/2 > 0$ for n sufficiently large in light of Assumption 6. This establishes (d). ■

Proof of Lemma A.5: Note from (1) and (10) that

$$\mathbf{y}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n}) = \mathbf{Z}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n})\boldsymbol{\delta}_{g,n} + \boldsymbol{\varepsilon}_{g,n} - (\widehat{\mathbf{R}}_{g,n}^{*} - \mathbf{R}_{g,n})\mathbf{u}_{g,n}$$

and hence

$$\begin{aligned}
& n^{1/2}[\widehat{\boldsymbol{\delta}}_{g,n} - \boldsymbol{\delta}_{g,n}] \\
= \; & \left[n^{-1}\widehat{\mathbf{Z}}_{*g,n}'(\widehat{\rho}_{g,n})\mathbf{Z}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n})\right]^{-1} n^{-1/2}\widehat{\mathbf{Z}}_{*g,n}'(\widehat{\boldsymbol{\rho}}_{g,n})\left[\boldsymbol{\varepsilon}_{g,n} - (\widehat{\mathbf{R}}_{g,n}^{*} - \mathbf{R}_{g,n})\mathbf{u}_{g,n}\right] \\
= \; & \widetilde{\mathbf{P}}_{gg,n}^{*\prime}\left[n^{-1/2}\mathbf{F}_{gg,n}^{*\prime}\boldsymbol{\varepsilon}_{g,n} - \sum_{r\in\mathbf{I}_{g,\rho}}(\widehat{\rho}_{g,r,n} - \rho_{g,r,n})n^{-1/2}\overline{\mathbf{F}}_{gg,r,n}^{*\prime}\boldsymbol{\varepsilon}_{g,n}\right],
\end{aligned}$$

with $\widehat{\mathbf{R}}_{g,n}^{*} = \mathbf{R}_{g,n}^{*}(\widehat{\boldsymbol{\rho}}_{g,n})$, and where $\widetilde{\mathbf{P}}_{gg,n}^{*}$ is defined in the lemma, $\mathbf{F}_{gg,n}^{*} = \mathbf{H}_n$, and $\overline{\mathbf{F}}_{gg,r,n}^{*} = \left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime}\right)^{-1}\mathbf{M}_{r,n}'\mathbf{H}_n$. In light of Assumption 6, and since $\widehat{\boldsymbol{\rho}}_n$ is consistent, it follows that

$$n^{-1}\widehat{\mathbf{Z}}_{*g,n}'(\widehat{\boldsymbol{\rho}}_{g,n})\mathbf{Z}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n}) - \mathbf{Q}_{\mathbf{HZ},g*}'(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n}) = o_p(1).$$

Since by Assumption 6 we have $\mathbf{Q}_{\mathbf{HZ},g*}'(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n}) = O(1)$ and

$[\mathbf{Q}'_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})]^{-1} = O(1)$ it follows that

$$[n^{-1}\widehat{\mathbf{Z}}'_{*g,n}(\widehat{\rho}_{g,n})\mathbf{Z}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n})]^{-1} - [\mathbf{Q}'_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})]^{-1} = o_p(1);$$

compare, e.g., Pötscher and Prucha (1997), Lemma F1. In light of this it follows further that $\widetilde{\mathbf{P}}^*_{gg,n} - \mathbf{P}^*_{gg,n} = o_p(1)$ and $\mathbf{P}^*_{gg,n} = O(1)$, with $\mathbf{P}^*_{gg,n}$ defined in the lemma. By argumentation analogous to that in the proof of Lemma A.4 it is readily seen that $n^{-1/2}\mathbf{F}^{*\prime}_{gg,n}\boldsymbol{\varepsilon}_{g,n} = O_p(1)$ and $n^{-1/2}\overline{\mathbf{F}}^{*\prime}_{gg,r,n}\boldsymbol{\varepsilon}_{g,n} = O_p(1)$. Consequently $n^{1/2}[\widehat{\boldsymbol{\delta}}_{g,n} - \boldsymbol{\delta}_{g,n}] = \mathbf{P}^{*\prime}_{gg,n}n^{-1/2}\mathbf{F}^{*\prime}_{gg,n}\boldsymbol{\varepsilon}_{g,n} + o_p(1)$ and $\mathbf{P}^{*\prime}_{gg,n}n^{-1/2}\mathbf{F}^{*\prime}_{gg,n}\varepsilon_{g,n} = O_p(1)$, observing again that $\widehat{\boldsymbol{\rho}}_{g,n} - \boldsymbol{\rho}_{g,n} = o_p(1)$. This established (a)-(c) recalling that $\mathbf{T}^*_{gg,n} = \mathbf{F}^*_{gg,n}\mathbf{P}^*_{gg,n}$. Next observe that

$$\begin{aligned}
& \lambda_{\min}(n^{-1}\mathbf{T}^{*\prime}_{gg,n}\mathbf{T}^*_{gg,n}) \\
\geq\ & \lambda_{\min}\left[\mathbf{Q}_{\mathbf{HH}}^{\prime -1/2}n^{-1}\mathbf{H}'_n\mathbf{H}_n\mathbf{Q}_{\mathbf{HH}}^{-1/2}\right] \\
& \lambda_{\min}\left\{[\mathbf{Q}'_{\mathbf{HZ},g*}(\rho_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})]^{-1}\right\} \\
\geq\ & \lambda_{\min}(\mathbf{Q}_{\mathbf{HH}}^{-1})\lambda_{\min}\left\{[\mathbf{Q}'_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})]^{-1}\right\}\lambda_{\min}\left[n^{-1}\mathbf{H}'_n\mathbf{H}_n\right] \\
\geq\ & \lambda_{\min}(\mathbf{Q}_{\mathbf{HH}}^{-1})\lambda_{\min}\left\{[\mathbf{Q}'_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})]^{-1}\right\}\left[\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}}^{-1})\right]/2 \geq c_*
\end{aligned}$$

for some $c_* > 0$ in light of Assumption 6, and observing that $\lambda_{\min}\left[n^{-1}\mathbf{H}'_n\mathbf{H}_n\right] \geq \left[\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}})\right]/2 > 0$ for n sufficiently large. This establishes (d). ■

Proof of Lemma A.6: Note from (1) and (11) that

$$\mathbf{y}_{*n}(\widehat{\boldsymbol{\rho}}_n) = \mathbf{Z}_{*n}(\widehat{\boldsymbol{\rho}}_n)\boldsymbol{\delta}_n + \boldsymbol{\varepsilon}_n - (\widehat{\mathbf{R}}^*_n - \mathbf{R}_n)\mathbf{u}_n$$

and hence

$$\begin{aligned}
& n^{1/2}[\widehat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_n] \\
=\ & \left[n^{-1}\widehat{\mathbf{Z}}'_{*n}(\widehat{\rho}_n)(\widehat{\boldsymbol{\Sigma}}_n^{-1}\otimes\mathbf{I}_n)\mathbf{Z}_{*n}(\widehat{\boldsymbol{\rho}}_n)\right]^{-1} n^{-1/2}\widehat{\mathbf{Z}}'_{*n}(\widehat{\boldsymbol{\rho}}_n)(\widehat{\boldsymbol{\Sigma}}_n^{-1}\otimes\mathbf{I}_n) \\
& \times\left[\boldsymbol{\varepsilon}_n - (\widehat{\mathbf{R}}^*_n - \mathbf{R}_n)\mathbf{u}_n\right] \\
=\ & \left[n^{-1}\widehat{\mathbf{Z}}'_{*n}(\widehat{\rho}_n)(\widehat{\boldsymbol{\Sigma}}_n^{-1}\otimes\mathbf{I}_n)\mathbf{Z}_{*n}(\widehat{\boldsymbol{\rho}}_n)\right]^{-1} diag\left[n^{-1}\mathbf{Z}'_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n})\mathbf{H}_n\right] \\
& \times\left[\widehat{\boldsymbol{\Sigma}}_n^{-1}\otimes(n^{-1}\mathbf{H}'_n\mathbf{H}_n)^{-1}\right](\mathbf{I}_G\otimes n^{-1/2}\mathbf{H}'_n)\left[\boldsymbol{\varepsilon}_n - (\widehat{\mathbf{R}}^*_n - \mathbf{R}_n)\mathbf{u}_n\right] \\
=\ & \widetilde{\mathbf{P}}^{**\prime}_n n^{-1/2}\mathbf{F}^{**\prime}_n\boldsymbol{\varepsilon}_n + \widetilde{\mathbf{P}}^{**\prime}_n\begin{bmatrix} \sum_{r\in\mathbf{I}_{1,\rho}}(\widehat{\rho}_{1,r,n} - \rho_{1,r,n})n^{-1/2}\overline{\mathbf{F}}^{**\prime}_{11,r,n}\varepsilon_{1,n} \\ \vdots \\ \sum_{r\in\mathbf{I}_{G,\rho}}(\widehat{\rho}_{G,r,n} - \rho_{G,r,n})n^{-1/2}\overline{\mathbf{F}}^{**\prime}_{GG,r,n}\varepsilon_{g,n}\end{bmatrix},
\end{aligned}$$

with $\widehat{\mathbf{R}}_n^* = diag_{g=1}^G\left[\mathbf{R}_{g,n}^*(\widehat{\boldsymbol{\rho}}_{g,n})\right]$, and where $\widetilde{\mathbf{P}}_n^{**}$ and $\mathbf{F}_n^{**}$ are defined in the lemma, and $\overline{\mathbf{F}}_{gg,r,n}^{**} = \left(\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime}\right)^{-1}\mathbf{M}_{r,n}^{\prime}\mathbf{H}_n$. Observe that the (g,h)-th block of

$$n^{-1}\widehat{\mathbf{Z}}_{*n}^{\prime}(\widehat{\rho}_n)(\widehat{\mathbf{\Sigma}}_n^{-1}\otimes\mathbf{I}_n)\mathbf{Z}_{*n}(\widehat{\rho}_n)-diag\left[\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\right]\left[\mathbf{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]$$

is

$$\widehat{\sigma}_{gh,n}n^{-1}\widehat{\mathbf{Z}}_{*g,n}^{\prime}(\widehat{\boldsymbol{\rho}}_{g,n})\mathbf{Z}_{*h,n}(\widehat{\boldsymbol{\rho}}_{h,n}) - \sigma_{gh}\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},h*}(\rho_{h,n}) = o_p(1)$$

in light of Assumption 6, and since $\widehat{\boldsymbol{\rho}}_n$ and $\widehat{\mathbf{\Sigma}}_n$ are consistent. By Assumption 6 we have $diag\left[\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\right]\left[\mathbf{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\rho_{g,n})\right] = O(1)$ and

$$\begin{aligned}&\lambda_{\min}\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\right]\left[\mathbf{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\}\\ \geq\ &\lambda_{\min}\left\{\mathbf{\Sigma}^{-1}\otimes\mathbf{I}\right\}\lambda_{\min}\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\mathbf{Q}_{\mathbf{HH}}^{-1}\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\} \geq c_*\end{aligned}$$

for some $c_* > 0$. This in turn implies that

$$\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\right]\left[\mathbf{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\}^{-1} = O(1).$$

Consequently

$$\begin{aligned}&\left\{n^{-1}\widehat{\mathbf{Z}}_{*n}^{\prime}(\widehat{\rho}_n)(\widehat{\mathbf{\Sigma}}_n^{-1}\otimes\mathbf{I}_n)\mathbf{Z}_{*n}(\widehat{\rho}_n)\right\}^{-1} - \\ &\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}^{\prime}(\boldsymbol{\rho}_{g,n})\right]\left[\mathbf{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\}^{-1} = o_p(1);\end{aligned}$$

compare, e.g., Pötscher and Prucha (1997), Lemma F1. In light of this it is now readily seen that $\widetilde{\mathbf{P}}_n^{**} - \mathbf{P}_n^{**} = o_p(1)$ and $\mathbf{P}_n^{**} = O(1)$, with $\mathbf{P}_n^{**}$ defined in the lemma. By argumentation analogous to that in the proof of Lemma A.4 it is readily seen that $n^{-1/2}\mathbf{F}_n^{**\prime}\boldsymbol{\varepsilon}_n = O_p(1)$ and $n^{-1/2}\overline{\mathbf{F}}_{gg,r,n}^{*\prime}\boldsymbol{\varepsilon}_{g,n} = O_p(1)$. Consequently $n^{1/2}[\widehat{\widehat{\boldsymbol{\delta}}}_n - \boldsymbol{\delta}_n] = \mathbf{P}_n^{**\prime}n^{-1/2}\mathbf{F}_n^{**\prime}\varepsilon_n + o_p(1)$ with $\mathbf{P}_n^{**\prime}n^{-1/2}\mathbf{F}_n^{**\prime}\varepsilon_n = O_p(1)$, observing again that $\widehat{\boldsymbol{\rho}}_{g,n} - \boldsymbol{\rho}_{g,n} = o_p(1)$. This establishes (a)-(c), recalling that

$\mathbf{T}_n^{**} = \mathbf{F}_n^{**}\mathbf{P}_n^{**}$. Next observe that

$$
\begin{aligned}
& \lambda_{\min}(n^{-1}\mathbf{T}_n^{**\prime}\mathbf{T}_n^{**}) \\
\geq \; & \lambda_{\min}(\Sigma^{-1})\lambda_{\min}\left[\mathbf{Q}_{\mathbf{HH}}^{\prime -1/2} n^{-1}\mathbf{H}_n'\mathbf{H}_n\mathbf{Q}_{\mathbf{HH}}^{-1/2}\right] \\
& \times\lambda_{\min}\left[\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}'(\boldsymbol{\rho}_{g,n})\right]\left[\boldsymbol{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\}^{-1}\right] \\
\geq \; & \lambda_{\min}(\Sigma^{-1})\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}}^{-1})\lambda_{\min}\left[n^{-1}\mathbf{H}_n'\mathbf{H}_n\right] \\
& \times\lambda_{\min}\left[\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}'(\boldsymbol{\rho}_{g,n})\right]\left[\boldsymbol{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\}^{-1}\right] \\
\geq \; & \lambda_{\min}(\Sigma^{-1})\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}}^{-1})[\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}}^{-1})/2] \\
& \times\lambda_{\min}\left[\left\{diag\left[\mathbf{Q}_{\mathbf{HZ},g*}'(\boldsymbol{\rho}_{g,n})\right]\left[\boldsymbol{\Sigma}^{-1}\otimes\mathbf{Q}_{\mathbf{HH}}^{-1}\right]diag\left[\mathbf{Q}_{\mathbf{HZ},g*}(\boldsymbol{\rho}_{g,n})\right]\right\}^{-1}\right] \geq c_*
\end{aligned}
$$

for some $c_* > 0$ in light of Assumption 6, and observing that $\lambda_{\min}\left[n^{-1}\mathbf{H}_n'\mathbf{H}_n\right] \geq \left[\lambda_{\min}(\mathbf{Q}_{\mathbf{HH}})\right]/2 > 0$ for n sufficiently large. This establishes (d). ■

Proof of Lemma A.7: Without loss of generality, assume that $\sigma^2 = 1$, since the model in Assumption A.1 can always be normalized accordingly.

We first prove part (a) of the lemma. Let $\vartheta_n = n^{-1}\mathfrak{u}_n'\mathbf{A}_n^*\mathfrak{u}_n$ and $\widetilde{\vartheta}_n = n^{-1}\widetilde{\mathfrak{u}}_n'\mathbf{A}_n^*\widetilde{\mathfrak{u}}_n$, then, in light of Assumption A.1, we have $\vartheta_n = n^{-1}\mathfrak{e}_n'\mathbf{B}_n^*\mathfrak{e}_n$ with $\mathbf{B}_n^* = (1/2)\mathfrak{R}_n^{-1\prime}(\mathbf{A}_n^* + \mathbf{A}_n^{*\prime})\mathfrak{R}_n^{-1}$. Furthermore, by Assumption A.1, the row and column sums of the matrices $\mathfrak{R}_n$ are uniformly bounded in absolute value. Since this property is preserved under matrix addition and multiplication - see, e.g., Remark A.1 in Kelejian and Prucha (2004) - it follows that also the row and column sums of the matrices $\mathbf{B}_n^*$ and $\mathbf{B}_n^*\mathbf{B}_n^*$ are uniformly bounded in absolute value. In the following let $K < \infty$ be a common bound for the row and column sums of the absolute elements of $\mathbf{B}_n^*$ and $\mathbf{B}_n^*\mathbf{B}_n^*$, and of their respective elements. Then, using the triangle inequality and the Cauchy-Schwarz inequality, we have

$$
E\left|\vartheta_n\right| = n^{-1}\sum_{i=1}^n\sum_{j=1}^n\left|b_{ij,n}^*\right|E\left|\mathfrak{e}_{i,n}\right|\left|\mathfrak{e}_{j,n}\right| \leq n^{-1}\sum_{i=1}^n\sum_{j=1}^n\left|b_{ij,n}^*\right| \leq K.
$$

Furthermore, utilizing the expression for the variance of linear quadratic forms given in Lemma A.1 in Kelejian and Prucha (2007) we have in light of Assumption A.1

$$
\begin{aligned}
var(\vartheta_n) &= n^{-2}2tr(\mathbf{B}_n^*\mathbf{B}_n^*) + n^{-2}\sum_{i=1}^n b_{ii,n}^{*2}\left[E\mathfrak{e}_{i,n}^4 - 3\right] \\
&\leq n^{-1}2K + n^{-1}K^2\sup_i\left[E\mathfrak{e}_{i,n}^4 - 3\right].
\end{aligned}
$$

Given that the fourth moments of the $\mathfrak{e}_{i,n}$ are uniformly bounded in light of Assumption A.1, this establishes the first two claims of part (a) of the lemma.

We next prove the last claim of part (a) of the lemma. The above discussion implies that $\vartheta_n - E\vartheta_n = o_p(1)$. Hence it suffices to show that $\widetilde{\vartheta}_n - \vartheta_n = o_p(1)$. By Assumptions A.1 and A.2

$$\widetilde{\vartheta}_n - \vartheta_n = \phi_n + \psi_n$$

with

$$\begin{aligned}\phi_n &= n^{-1}\mathbf{\Delta}_n'\mathfrak{D}_n'\left(\mathbf{A}_n^* + \mathbf{A}_n^{*\prime}\right)\mathfrak{u}_n = n^{-1}\mathbf{\Delta}_n'\mathfrak{D}_n'\mathbf{C}_n^*\mathfrak{e}_n, \\ \psi_n &= n^{-1}\mathbf{\Delta}_n'\mathfrak{D}_n'\mathbf{A}_n^*\mathfrak{D}_n\mathbf{\Delta}_n,\end{aligned}$$

and $\mathbf{C}_n^* = (c_{ij,n}^*) = \left(\mathbf{A}_n^* + \mathbf{A}_n^{*\prime}\right)\mathfrak{R}_n^{-1}$. The row and column sums of the matrices $\mathbf{C}_n^*$ are again seen to be uniformly bounded in absolute value. Let $\overline{K} < \infty$ denote a uniform bound for the row and column sums of the absolute elements of the matrices $\mathbf{A}_n^*$ and $\mathbf{C}_n^*$, and let $\mathbf{c}_{i.,n}^*$ and $\mathfrak{d}_{i.,n}$ denote the i-th row of $\mathbf{C}_n^*$ and $\mathfrak{D}_n$, respectively.

To prove the claim we now show that both ϕ_n and ψ_n are $o_p(1)$. Using the triangle and Hölder inequality we get

$$\begin{aligned}|\phi_n| &= \left| n^{-1}\sum_{i=1}^{n}\mathbf{\Delta}_n'\mathfrak{d}_{i.,n}'\mathbf{c}_{i.,n}^*\mathfrak{e}_n \right| \qquad \text{(D.2)} \\ &\leq n^{-1}\left\|\mathbf{\Delta}_n\right\|\sum_{i=1}^{n}\left\|\mathfrak{d}_{i.,n}\right\|\sum_{j=1}^{n}\left|c_{ij,n}^*\right|\left|\mathfrak{e}_{j,n}\right| \leq n^{-1}\left\|\mathbf{\Delta}_n\right\|\sum_{j=1}^{n}\left|\mathfrak{e}_{j,n}\right|\sum_{i=1}^{n}\left\|\mathfrak{d}_{i.,n}\right\|\left|c_{ij,n}^*\right| \\ &\leq n^{-1}\left\|\mathbf{\Delta}_n\right\|\sum_{j=1}^{n}\left|\mathfrak{e}_{j,n}\right|\left(\sum_{i=1}^{n}\left\|\mathfrak{d}_{i.,n}\right\|^p\right)^{1/p}\left(\sum_{i=1}^{n}\left|c_{ij,n}^*\right|^q\right)^{1/q} \\ &\leq \overline{K}n^{1/p-1/2}\left(n^{1/2}\left\|\mathbf{\Delta}_n\right\|\right)\left(n^{-1}\sum_{j=1}^{n}\left|\mathfrak{e}_{j,n}\right|\right)\left(n^{-1}\sum_{i=1}^{n}\left\|\mathfrak{d}_{i.,n}\right\|^p\right)^{1/p}\end{aligned}$$

for $p = 2+\delta$ and $1/p + 1/q = 1$, and where $\delta > 0$ is as in Assumption A.2. The last inequality utilizes the observation of Remark C.1 in Kelejian and Prucha (2007). Since the $\mathfrak{e}_{j,n}$ are independent with bounded second moments, it follows that $n^{-1}\sum_{j=1}^{n}\left|\mathfrak{e}_{j,n}\right| = O_p(1)$. The terms $n^{1/2}\left\|\mathbf{\Delta}_n\right\|$ and $n^{-1}\sum_{i=1}^{n}\left\|\mathfrak{d}_{i.,n}\right\|^p$ are $O_p(1)$ by Assumption A.2. Since $n^{1/p-1/2} \to 0$ as $n \to \infty$ it follows that $\phi_n = o_p(1)$.

Again, using the triangle and Hölder inequality yields

$$\begin{aligned} |\psi_n| &= \left| n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{\Delta}_n' \mathfrak{d}_{i.,n}' a_{ij,n}^* \mathfrak{d}_{j.,n} \mathbf{\Delta}_n \right| \qquad \text{(D.3)} \\ &\leq n^{-1} \|\mathbf{\Delta}_n\|^2 \sum_{i=1}^{n} \|\mathfrak{d}_{i.,n}\| \sum_{j=1}^{n} \|\mathfrak{d}_{j.,n}\| \left|a_{ij,n}^*\right| \\ &\leq n^{-1} \|\mathbf{\Delta}_n\|^2 \sum_{i=1}^{n} \|\mathfrak{d}_{i.,n}\| \left(\sum_{j=1}^{n} \|\mathfrak{d}_{j.,n}\|^p \right)^{1/p} \left(\sum_{j=1}^{n} \left|a_{ij,n}^*\right|^q \right)^{1/q} \\ &\leq \overline{K} n^{1/p} \|\mathbf{\Delta}_n\|^2 \left(n^{-1} \sum_{i=1}^{n} \|\mathfrak{d}_{i.,n}\| \right) \left(n^{-1} \sum_{j=1}^{n} \|\mathfrak{d}_{j.,n}\|^p \right)^{1/p} \\ &\leq \overline{K} n^{1/p-1/2} n^{-1/2} (n^{1/2} \|\mathbf{\Delta}_n\|)^2 \left(n^{-1} \sum_{i=1}^{n} \|\mathfrak{d}_{i.,n}\|^p \right)^{2/p} \end{aligned}$$

with p and q as before. By Assumption A.2 both $n^{-1}\sum_{i=1}^{n} \|\mathfrak{d}_{i.,n}\|^p$ and $n^{1/2}\|\mathbf{\Delta}_n\|$ are $O_p(1)$. Since $n^{1/p-1/2} \to 0$ as $n \to \infty$ it follows that $\psi_n = o_p(1)$. From the last inequality we see also that $n^{1/2}\psi_n = o_p(1)$.

We next prove part (b) of the lemma. In the following let $\underline{\vartheta}_{s,n}$ denote the s-th element of $n^{-1}\mathfrak{D}_n'\mathbf{A}_n^*\mathfrak{u}_n$. Observe $E\mathfrak{u}_n\mathfrak{u}_n' = \mathfrak{R}_n^{-1}\mathfrak{R}_n^{-1\prime}$. Then given Assumptions A.1 and A.2 there exists a constant $\overline{\overline{K}} < \infty$ such that $E\mathfrak{u}_{i,n}^2 \leq \overline{\overline{K}}$ and $E|\mathfrak{d}_{js,n}|^p \leq \overline{\overline{K}}$. WLOG assume that the row and column sums of the matrices $\mathbf{A}_n^*$ are uniformly bounded by $\overline{\overline{K}}$. Utilizing the Cauchy-Schwarz and Lyapunov inequalities we then have $E|\mathfrak{u}_{i,n}||\mathfrak{d}_{js,n}| \leq \left[E\mathfrak{u}_{i,n}^2\right]^{1/2}\left[E\mathfrak{d}_{js,n}^2\right]^{1/2} \leq \left[E\mathfrak{u}_{i,n}^2\right]^{1/2}(E|\mathfrak{d}_{js,n}|^p)^{1/p} \leq \overline{\overline{K}}^{1/2+1/p}$ with p as before and, hence,

$$E\left|\underline{\vartheta}_{s,n}\right| = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij,n}^*\right| E\left[|\mathfrak{u}_{i,n}||\mathfrak{d}_{js,n}|\right] \leq \overline{\overline{K}}^{1/2+1/p} n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij,n}^*\right| \leq \overline{\overline{K}}^{3/2+1/p} < \infty,$$

which shows that indeed $E\left|n^{-1}\mathfrak{d}_{.s,n}'\mathbf{A}_n^*\mathfrak{u}_n\right| = O(1)$ where $\mathfrak{d}_{.s,n}$ denotes the s-th column of $\mathfrak{D}_n$. Of course, the argument also shows that $\boldsymbol{\alpha}_n^* = n^{-1}E\mathfrak{D}_n'(\mathbf{A}_n^* + \mathbf{A}_n^{*\prime})\mathfrak{u}_n = O(1)$. Next observe that

$$n^{-1}\mathfrak{D}_n'\mathbf{A}_n^*\widetilde{\mathfrak{u}}_n = n^{-1}\mathfrak{D}_n'\mathbf{A}_n^*\mathfrak{u}_n + \underline{\psi}_n,$$

where $\underline{\psi}_n = n^{-1}\mathfrak{D}_n'\mathbf{A}_n^*\mathfrak{D}_n\mathbf{\Delta}_n$. By argumentation analogous to that employed to demonstrate that $n^{1/2}\psi_n = o_p(1)$ it follows that also $\underline{\psi}_n = o_p(1)$, which completes the proof of part (b).

We next prove part (c). In light of the proof of part (a) we have

$$n^{-1/2}\widetilde{\mathfrak{u}}_n'\mathbf{A}_n^*\widetilde{\mathfrak{u}}_n = n^{-1/2}\mathfrak{u}_n'\mathbf{A}_n^*\mathfrak{u}_n + [n^{-1}\mathfrak{u}_n'(\mathbf{A}_n^* + \mathbf{A}_n^{*\prime})\mathfrak{D}_n]n^{1/2}\mathbf{\Delta}_n + n^{1/2}\psi_n$$

with $n^{1/2}\psi_n = o_p(1)$. In light of part (b) and Assumption A.3 we have $n^{-1}\mathfrak{u}_n'(\mathbf{A}_n^* + \mathbf{A}_n^{*\prime})\mathfrak{D}_n - \boldsymbol{\alpha}_n^{*\prime} = o_p(1)$. The claim follows since $n^{1/2}\mathbf{\Delta}_n = O_p(1)$ by Assumption A.2. ■

Remark A.1: For future reference it proves helpful to note that in light of Remark A.1 in Kelejian and Prucha (2004) the constant $\overline{K}$ used in proving the last claim of part (a) of the above lemma can be chosen as $\overline{K} = 2c_P c_A$ where c_P and c_A denote a bound for the row and column sums of the absolute elements of $\mathfrak{R}_n^{*-1}$ and $\mathbf{A}_n^*$. Furthermore it proves helpful to observe that in light of (D.2) and (D.3)

$$\left|\widetilde{\vartheta}_n - \vartheta_n\right| \leq 2c_P c_A \varsigma_n,$$

where $\varsigma_n = o_p(1)$ does not depend on $\mathbf{A}_n^*$.

Proof of Lemma A.8: Given Assumption A.1 and the maintained assumptions on $\mathbf{A}_n$ it follows that the row and column sums of $\mathbf{A}_n^* = \mathfrak{R}_n'\mathbf{A}_n\mathfrak{R}_n$ are bounded uniformly in absolute value. Thus by Lemma A.7(c), and utilizing Assumption A.4, we have

$$\begin{aligned}
& n^{-1/2}\widetilde{\mathfrak{u}}_n'\mathfrak{R}_n'\mathbf{A}_n\mathfrak{R}_n\widetilde{\mathfrak{u}}_n \\
= \;& n^{-1/2}\mathfrak{u}_n'\mathfrak{R}_n'\mathbf{A}_n\mathfrak{R}_n\mathfrak{u}_n + \boldsymbol{\alpha}_n' n^{1/2}\Delta_n + o_p(1) \\
= \;& n^{-1/2}\boldsymbol{\varepsilon}_{g,n}'\mathbf{A}_n\varepsilon_{g,n} + n^{-1/2}\boldsymbol{\alpha}_n'\left[\sum_{h=1}^{G}\mathbf{T}_{h,n}'\boldsymbol{\varepsilon}_{h,n} + o_p(1)\right] + o_p(1) \\
= \;& n^{-1/2}\boldsymbol{\varepsilon}_{g,n}'\mathbf{A}_n\varepsilon_{g,n} + n^{-1/2}\sum_{h=1}^{G}\mathbf{a}_{h,n}'\varepsilon_{h,n} + o_p(1).
\end{aligned}$$

The last inequality holds since $\boldsymbol{\alpha}_n = O(1)$; see the remark in Lemma A.7(c). Given this and the maintained assumption on $\mathbf{P}_{h,n}$ it follows that $\mathbf{c}_{h,n} = (c_{h1,n}, \ldots, c_{hp_\Delta,n})' = \mathbf{P}_{h,n}\boldsymbol{\alpha}_n = O(1)$. Since $\mathbf{a}_{h,n} = \mathbf{F}_{h,n}\mathbf{c}_{h,n}$ we have

$$|a_{hi,n}|^\eta = \left|\sum_{s=1}^{p_F} f_{his,n}c_{hs,n}\right|^\eta \leq p_F^\eta K^\eta \sum_{s=1}^{p_\Delta}|f_{his,n}|^\eta,$$

using inequality (1.4.4.) in Bierens (1994). Thus

$$\sup_n n^{-1} \sum_{i=1}^{n} |a_{hi,n}|^\eta \leq p_F^\eta K^\eta \sum_{s=1}^{pF} \sup_n n^{-1} \sum_{i=1}^{n} |f_{his,n}|^\eta < \infty$$

in light of Assumption A.4(a). This proves part (a). Part (b) follows readily from, e.g., Lemma A.1 in Kelejian and Prucha (2010). ■

E Supplement to Section 6

In the following we provide more details on the derivation of the results in the Section on Limited and Full Information One-Step Estimators. We first discuss limited information estimators. Let $\boldsymbol{\theta}_{g,n} = [\delta'_{g,n}, \boldsymbol{\rho}'_{g,n}]'$ and $\widehat{\boldsymbol{\theta}}^{o}_{g,n} = [\widehat{\boldsymbol{\delta}}^{o\prime}_{g,n}, \widehat{\boldsymbol{\rho}}^{o\prime}_{g,n}]'$, then by argumentation analogous to those for two-step estimators we have $(\boldsymbol{\Phi}^{o}_{gg,n})^{-1/2} n^{1/2}(\widehat{\boldsymbol{\theta}}^{o}_{g,n} - \boldsymbol{\theta}_{g,n}) \xrightarrow{d} N[0, I]$ with

$$\boldsymbol{\Phi}^{o}_{gg,n} - \left[\left[\frac{\partial \mathbf{m}_{g,n}(\boldsymbol{\theta}_{g,n})}{\partial \boldsymbol{\theta}'_{g,n}}\right] \begin{bmatrix} \boldsymbol{\Psi}^{LL}_{gg,n} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^{QQ}_{gg,n} \end{bmatrix}^{-1} \left[\frac{\partial \mathbf{m}_{g,n}(\boldsymbol{\theta}_{g,n})}{\partial \boldsymbol{\theta}_{g,n}}\right]\right]^{-1} = o_p(1),$$

where

$$\begin{aligned} \boldsymbol{\Psi}^{LL}_{gg,n} &= \sigma_{gg,n}\left[n^{-1}\mathbf{H}'_n \mathbf{H}_n\right], \\ \boldsymbol{\Psi}^{QQ}_{gg,n} &= \sigma^2_{gg,n}\mathbf{K}^{QQ}_n, \end{aligned}$$

with $\mathbf{K}^{QQ}_n = (k^{QQ}_{rs,n})$ and

$$k^{QQ}_{rs,n} = (2n)^{-1} tr\left[(\mathbf{A}_{r,n} + \mathbf{A}'_{r,n})(\mathbf{A}_{s,n} + \mathbf{A}'_{s,n})\right].$$

Recall that

$$\begin{aligned} \mathbf{y}_{g,n} &= \mathbf{Z}_{g,n}\boldsymbol{\delta}_{g,n} + \mathbf{u}_{g,n}, \\ \mathbf{u}_{g,n} &= \overline{\mathbf{U}}_{g,n}\boldsymbol{\rho}_{g,n} + \varepsilon_{g,n} = \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})\mathbf{u}_{g,n} + \boldsymbol{\varepsilon}_{g,n}, \end{aligned}$$

with $\mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n}) = \sum_{r \in \mathbf{I}_{g,\rho}} \rho_{g,r,n}\mathbf{M}_{r,n}$. Consequently

$$\begin{aligned} \boldsymbol{\varepsilon}_{g,n} &= [\mathbf{I}_n - \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})]\mathbf{u}_{g,n} \\ &= [\mathbf{I}_n - \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})][\mathbf{y}_{g,n} - \mathbf{Z}_{g,n}\boldsymbol{\delta}_{g,n}] \end{aligned}$$

and

$$\begin{aligned} \mathbf{H}'_n\boldsymbol{\varepsilon}_{g,n} &= \mathbf{H}'_n[\mathbf{I}_n - \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})][\mathbf{y}_{g,n} - \mathbf{Z}_{g,n}\boldsymbol{\delta}_{g,n}], \\ \boldsymbol{\varepsilon}'_{g,n}\mathbf{A}_{s,n}\boldsymbol{\varepsilon}_{g,n} &= \mathbf{u}'_{g,n}[\mathbf{I}_n - \mathbf{R}^{*\prime}_{g,n}(\boldsymbol{\rho}_{g,n})]\overline{\mathbf{A}}_{s,n}[\mathbf{I}_n - \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})]\mathbf{u}_{g,n} \\ &= \mathbf{u}'_{g,n}\overline{\mathbf{A}}_{s,n}\mathbf{u}_{g,n} - 2\mathbf{u}'_{g,n}\mathbf{R}^{*\prime}_{g,n}(\boldsymbol{\rho}_{g,n})\overline{\mathbf{A}}_{s,n}\mathbf{u}_{g,n} \\ &\quad + \mathbf{u}'_{g,n}\mathbf{R}^{*\prime}_{g,n}(\boldsymbol{\rho}_{g,n})\overline{\mathbf{A}}_{s,n}\mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})\mathbf{u}_{g,n}. \end{aligned}$$

Thus

$$\frac{\partial}{\partial \boldsymbol{\delta}_{g,n}}[\mathbf{H}_n'\boldsymbol{\varepsilon}_{g,n}] = \mathbf{H}_n'[\mathbf{I}_n - \mathbf{R}_{g,n}^*(\boldsymbol{\rho}_{g,n})]\frac{\partial \mathbf{u}_{g,n}}{\partial \boldsymbol{\delta}_{g,n}} = -\mathbf{H}_n'[\mathbf{I}_n - \mathbf{R}_{g,n}^*(\boldsymbol{\rho}_{g,n})]\mathbf{Z}_{g,n},$$

$$\begin{aligned} &\frac{\partial}{\partial \boldsymbol{\delta}_{g,n}}[\boldsymbol{\varepsilon}_{g,n}'\mathbf{A}_{s,n}\boldsymbol{\varepsilon}_{g,n}] = 2\mathbf{u}_{g,n}'[\mathbf{I}_n - \mathbf{R}_{g,n}^{*\prime}(\boldsymbol{\rho}_{g,n})]\overline{\mathbf{A}}_{s,n}[\mathbf{I}_n - \mathbf{R}_{g,n}^*(\boldsymbol{\rho}_{g,n})]\frac{\partial \mathbf{u}_{g,n}}{\partial \delta_{g,n}} \\ = \quad & -2\boldsymbol{\varepsilon}_{g,n}'\overline{\mathbf{A}}_{s,n}[\mathbf{I}_n - \mathbf{R}_{g,n}^*(\boldsymbol{\rho}_{g,n})]\mathbf{Z}_{g,n} \\ = \quad & -\boldsymbol{\varepsilon}_{g,n}'(\mathbf{A}_{s,n} + \mathbf{A}_{s,n}')[\mathbf{I}_n - \mathbf{R}_{g,n}^*(\boldsymbol{\rho}_{g,n})]\mathbf{Z}_{g,n}, \end{aligned}$$

and

$$\frac{\partial}{\partial \boldsymbol{\rho}_{g,n}}[\mathbf{H}_n'\boldsymbol{\varepsilon}_{g,n}] = -[\mathbf{H}_n'\mathbf{M}_{r_{g,1},n}\mathbf{u}_{g,n}, \ldots, \mathbf{H}_n'\mathbf{M}_{r_{g,q_g},n}\mathbf{u}_{g,n}],$$

$$\begin{aligned} &\frac{\partial}{\partial \boldsymbol{\rho}_{g,n}}[\boldsymbol{\varepsilon}_{g,n}'\mathbf{A}_{s,n}\boldsymbol{\varepsilon}_{g,n}] = -2[\mathbf{u}_{g,n}'\mathbf{M}_{r_{g,1},n}'\overline{\mathbf{A}}_{s,n}\mathbf{u}_{g,n}, ..., \mathbf{u}_{g,n}'\mathbf{M}_{r_{g,q_g},n}'\overline{\mathbf{A}}_{s,n}\mathbf{u}_{g,n})]\frac{\partial \mathbf{r}_{1,g,n}}{\partial \boldsymbol{\rho}_{g,n}} \\ &+[\mathbf{u}_{g,n}'\mathbf{M}_{r_{g,1},n}'\overline{\mathbf{A}}_{s,n}\mathbf{M}_{r_{g,1},n}\mathbf{u}_{g,n}, ..., \mathbf{u}_{g,n}'\mathbf{M}_{r_{g,q_g},n}'\overline{\mathbf{A}}_{s,n}\mathbf{M}_{r_{g,q_g},n}\mathbf{u}_{g_g,n}]\frac{\partial \mathbf{r}_{2,g,n}}{\partial \boldsymbol{\rho}_{g,n}} \\ &+2[E\mathbf{u}_{g,n}'\mathbf{M}_{r_{g,1},n}'\overline{\mathbf{A}}_{s,n}\mathbf{M}_{r_{g,2},n}\mathbf{u}_{g,n}, ..., E\mathbf{u}_{g,n}'\mathbf{M}_{r_{g,q_g-1},n}'\overline{\mathbf{A}}_{s,n}\mathbf{M}_{r_{g,q_g},n}\mathbf{u}_{g,n}]\frac{\partial \mathbf{r}_{3,g,n}}{\partial \boldsymbol{\rho}_{g,n}}. \end{aligned}$$

Recalling that

$$\begin{aligned} \mathbf{m}_{g,n}^{\delta}(\boldsymbol{\rho}_{g,n}, \boldsymbol{\delta}_{g,n}) &= n^{-1}\mathbf{H}_n'\boldsymbol{\varepsilon}_{g,n}, \\ \mathbf{m}_{g,n}^{\rho}(\boldsymbol{\rho}_{g,n}, \boldsymbol{\delta}_{g,n}) &= \begin{bmatrix} n^{-1}\boldsymbol{\varepsilon}_{g,n}'\mathbf{A}_{1,n}\varepsilon_{g,n} \\ \vdots \\ n^{-1}\boldsymbol{\varepsilon}_{g,n}'\mathbf{A}_{S,n}\varepsilon_{g,n} \end{bmatrix}, \end{aligned}$$

we have

$$\frac{\partial \mathbf{m}^{\delta}_{g,n}}{\partial \boldsymbol{\delta}_{g,n}} = -n^{-1}\mathbf{H}'_n\mathbf{Z}_{*g,n}(\boldsymbol{\rho}_{g,n}) = -n^{-1}\left[\mathbf{H}'_n\mathbf{Z}_{g,n} - \sum_{r\in\mathbf{I}_{g,\rho}} \rho_{g,r,n}\mathbf{H}'_n\mathbf{M}_{r,n}\mathbf{Z}_{g,n}\right]$$
$$= -\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n}) + o_p(1),$$

$$\frac{\partial \mathbf{m}^{\delta}_{g,n}}{\partial \boldsymbol{\rho}_{g,n}} = -n^{-1}[\mathbf{H}'_n\mathbf{M}_{r_{g,1},n}\mathbf{u}_{g,n}, \ldots, \mathbf{H}'_n\mathbf{M}_{r_{g,q_g},n}\mathbf{u}_{g,n}] = o_p(1),$$

$$\frac{\partial \mathbf{m}^{\rho}_{g,n}}{\partial \boldsymbol{\delta}_{g,n}} = \begin{bmatrix} -n^{-1}\boldsymbol{\varepsilon}'_{g,n}[\mathbf{A}_{1,n}+\mathbf{A}'_{1,n}][\mathbf{I}_n - \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})]\mathbf{Z}_{g,n} \\ \vdots \\ -n^{-1}\boldsymbol{\varepsilon}'_{g,n}[\mathbf{A}_{S,n}+\mathbf{A}'_{S,n}][\mathbf{I}_n - \mathbf{R}^*_{g,n}(\boldsymbol{\rho}_{g,n})]\mathbf{Z}_{g,n} \end{bmatrix} = \begin{bmatrix} \alpha'_{g,1,n} \\ \vdots \\ \alpha'_{g,S,n} \end{bmatrix} + o_p(1),$$

$$\frac{\partial \mathbf{m}^{\rho}_{g,n}}{\partial \boldsymbol{\rho}_{g,n}} = -\boldsymbol{\Gamma}_{g,n}(\boldsymbol{\delta}_{g,n})\frac{\partial \mathbf{r}_{g,n}(\boldsymbol{\rho}_{g,n})}{\partial \boldsymbol{\rho}_{g,n}} + o_p(1) = -\mathbf{J}_{g,n} + o_p(1),$$

with $\boldsymbol{\alpha}_{g,r,n} = -n^{-1}E\mathbf{Z}'_{g,n}(\mathbf{I}_n - \mathbf{R}^{*\prime}_{g,n}(\boldsymbol{\rho}_{g,n}))(\mathbf{A}_{r,n}+\mathbf{A}'_{r,n})\boldsymbol{\varepsilon}_{g,n}$ and $\mathbf{J}_{g,n} = \boldsymbol{\Gamma}_{g,n}\frac{\partial \mathbf{r}_{g,n}(\boldsymbol{\rho}_{g,n})}{\partial \boldsymbol{\rho}_{g,n}}$. Consequently, we have

$$\frac{\partial \mathbf{m}_{g,n}}{\partial \boldsymbol{\theta}_{g,n}} = \mathbf{G}_{g,n} + o_p(1)$$

with

$$\mathbf{G}_{g,n} = \begin{bmatrix} \mathbf{G}^{LL}_{g,n} & \mathbf{G}^{LQ}_{g,n} \\ \mathbf{G}^{QL}_{g,n} & \mathbf{G}^{QQ}_{g,n} \end{bmatrix} = \begin{bmatrix} -\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n}) & \mathbf{0} \\ \boldsymbol{\alpha}'_{g,n} & -\mathbf{J}_{g,n} \end{bmatrix}$$

where $\boldsymbol{\alpha}_{g,n} = [\alpha_{g,1,n}, \ldots, \alpha_{g,S,n}]$. Furthermore

$$\mathbf{S}_{g,n} = \mathbf{G}'_{g,n}\begin{bmatrix} (\boldsymbol{\Psi}^{LL}_{gg,n})^{-1} & \mathbf{0} \\ \mathbf{0} & (\boldsymbol{\Psi}^{QQ}_{gg,n})^{-1} \end{bmatrix}\mathbf{G}_{g,n} = \begin{bmatrix} \mathbf{S}^{LL}_{g,n} & \mathbf{S}^{LQ}_{g,n} \\ \mathbf{S}^{QL}_{g,n} & \mathbf{S}^{QQ}_{g,n} \end{bmatrix}$$

where

$$\mathbf{S}^{LL}_{g,n} = \mathbf{G}^{LL\prime}_{g,n}(\boldsymbol{\Psi}^{LL}_{gg,n})^{-1}\mathbf{G}^{LL}_{g,n} + \mathbf{G}^{QL\prime}_{g,n}(\boldsymbol{\Psi}^{QQ}_{gg,n})^{-1}\mathbf{G}^{QL}_{g,n}$$
$$= \sigma^{-1}_{gg,n}\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n})'(n^{-1}\mathbf{H}'_n\mathbf{H}_n)^{-1}\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n}) + \sigma^{-2}_{gg,n}\alpha_{g,n}(\mathbf{K}^{QQ}_n)^{-1}\boldsymbol{\alpha}'_{g,n},$$
$$\mathbf{S}^{LQ}_{g,n} = \mathbf{S}^{LQ\prime}_{g,n} = \mathbf{G}^{QL\prime}_{g,n}(\boldsymbol{\Psi}^{QQ}_{gg,n})^{-1}\mathbf{G}^{QQ}_{g,n} = -\sigma^{-2}_{gg,n}\boldsymbol{\alpha}_{g,n}(\mathbf{K}^{QQ}_n)^{-1}\mathbf{J}_{g,n},$$
$$\mathbf{S}^{QQ}_{g,n} = \mathbf{G}^{QQ\prime}_{g,n}(\boldsymbol{\Psi}^{QQ}_{gg,n})^{-1}\mathbf{G}^{QQ}_{g,n} = \sigma^{-2}_{gg,n}\mathbf{J}'_{g,n}(\mathbf{K}^{QQ}_n)^{-1}\mathbf{J}_{g,n}.$$

The submatrices of $\mathbf{S}_{g,n}$ can be estimated consistently by

$$\begin{aligned}
\widehat{\mathbf{S}}_{g,n}^{o,LL} &= \widehat{\sigma}_{gg}^{-1}[n^{-1}\widehat{\mathbf{Z}}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n}^{o})'\widehat{\mathbf{Z}}_{*g,n}(\widehat{\boldsymbol{\rho}}_{g,n}^{o})] + \widehat{\sigma}_{gg,n}^{-2}\widehat{\boldsymbol{\alpha}}_{g,n}^{o}(\mathbf{K}_{n}^{QQ})^{-1}\widehat{\boldsymbol{\alpha}}_{g,n}^{o\prime}, \\
\widehat{\mathbf{S}}_{g,n}^{o,LQ} &= \widehat{\mathbf{S}}_{g,n}^{o,LQ\prime} = -\widehat{\sigma}_{gg,n}^{-2}\widehat{\boldsymbol{\alpha}}_{g,n}^{o}(\mathbf{K}_{n}^{QQ})^{-1}\mathbf{J}_{g,n}(\widehat{\boldsymbol{\rho}}_{g,n}^{o}), \\
\widehat{\mathbf{S}}_{g,n}^{o,QQ} &= \widehat{\sigma}_{gg,n}^{-2}\mathbf{J}_{g,n}'(\widehat{\boldsymbol{\rho}}_{g,n}^{o})\left(\mathbf{K}_{n}^{QQ}\right)^{-1}\mathbf{J}_{g,n}(\widehat{\boldsymbol{\rho}}_{g,n}^{o}),
\end{aligned}$$

where $\widehat{\boldsymbol{\alpha}}_{g,n}^{o} = [\widehat{\alpha}_{g,1,n}^{o}, \ldots, \widehat{\boldsymbol{\alpha}}_{g,S,n}^{o}]$ with

$$\widehat{\boldsymbol{\alpha}}_{g,r,n}^{o} = -n^{-1}\left[\mathbf{Z}_{*g,n}'(\widehat{\boldsymbol{\rho}}_{g,n}^{o})(\mathbf{A}_{r,n}+\mathbf{A}_{r,n}')(\mathbf{I}_{n}-\mathbf{R}_{g,n}^{*}(\widehat{\rho}_{g,n}^{o}))\widehat{\mathbf{u}}_{g,n}\right].$$

We next discuss full information estimators. Let $\boldsymbol{\theta}_n = [\boldsymbol{\delta}_n', \boldsymbol{\rho}_n']'$ and $\widehat{\widehat{\boldsymbol{\theta}}}_n^{o} = [\widehat{\widehat{\delta}}_n^{o\prime}, \widehat{\widehat{\boldsymbol{\rho}}}_n^{o\prime}]'$, then by argumentation analogous to those for two-step estimators we have $(\boldsymbol{\Phi}_n^{o})^{-1/2}n^{1/2}(\widehat{\widehat{\boldsymbol{\theta}}}_n^{o} - \boldsymbol{\theta}_n) \xrightarrow{d} N[0,I]$ with

$$\boldsymbol{\Phi}_n^{o} - \left[\left[\frac{\partial \mathbf{m}_n(\theta_n)}{\partial \boldsymbol{\theta}_n'}\right]\begin{bmatrix}\boldsymbol{\Psi}_n^{LL} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_n^{QQ}\end{bmatrix}^{-1}\left[\frac{\partial \mathbf{m}_n(\boldsymbol{\theta}_n)}{\partial \boldsymbol{\theta}_n}\right]\right]^{-1} = o_p(1),$$

where

$$\begin{aligned}
\boldsymbol{\Psi}_n^{LL} &= \boldsymbol{\Sigma}_n \otimes \left[n^{-1}\mathbf{H}_n'\mathbf{H}_n\right], \\
\boldsymbol{\Psi}_n^{QQ} &= \begin{bmatrix}\boldsymbol{\Psi}_{11,n}^{QQ} & \cdots & \boldsymbol{\Psi}_{1G,n}^{QQ} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{G1,n}^{QQ} & \cdots & \boldsymbol{\Psi}_{GG,n}^{QQ}\end{bmatrix} = \boldsymbol{\Sigma}_{SQ,n} \otimes \mathbf{K}_n^{QQ},
\end{aligned}$$

with $\boldsymbol{\Psi}_{gh,n}^{QQ} = \sigma_{gh,n}^2\mathbf{K}_n^{QQ}$ and $\boldsymbol{\Sigma}_{SQ,n} = (\sigma_{gh,n}^2)$.

Observe that for $g \neq h$,

$$\frac{\partial \mathbf{m}_{g,n}}{\partial \boldsymbol{\theta}_{h,n}} = 0,$$

and thus

$$\frac{\partial \mathbf{m}_n}{\partial \boldsymbol{\theta}_n} = \mathbf{G}_n + o_p(1)$$

with

$$\begin{aligned}
\mathbf{G}_n &= \begin{bmatrix} diag_{g=1}^{G}[\mathbf{G}_{g,n}^{LL}] & diag_{g=1}^{G}[\mathbf{G}_{g,n}^{LQ}] \\ diag_{g=1}^{G}[\mathbf{G}_{g,n}^{QL}] & diag_{g=1}^{G}[\mathbf{G}_{g,n}^{QQ}]\end{bmatrix} \\
&= \begin{bmatrix} -diag_{g=1}^{G}[\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n})] & \mathbf{0} \\ diag_{g=1}^{G}[\boldsymbol{\alpha}_{g,n}'] & -diag_{g=1}^{G}[\mathbf{J}_{g,n}]\end{bmatrix}.
\end{aligned}$$

Furthermore

$$\mathbf{S}_n = \mathbf{G}_n' \begin{bmatrix} (\boldsymbol{\Psi}_n^{LL})^{-1} & \mathbf{0} \\ \mathbf{0} & (\boldsymbol{\Psi}_n^{QQ})^{-1} \end{bmatrix} \mathbf{G}_n = \begin{bmatrix} \mathbf{S}_n^{LL} & \mathbf{S}_n^{LQ} \\ \mathbf{S}_n^{QL} & \mathbf{S}_n^{QQ} \end{bmatrix}$$

where

$$\begin{aligned}
\mathbf{S}_n^{LL} &= \mathbf{G}_n^{LL\prime}(\boldsymbol{\Psi}_n^{LL})^{-1}\mathbf{G}_n^{LL} + \mathbf{G}_n^{QL\prime}(\boldsymbol{\Psi}_n^{QQ})^{-1}\mathbf{G}_n^{QL} \\
&= diag_{g=1}^{G}[\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n})'][\boldsymbol{\Sigma}_{SQ,n}^{-1} \otimes n^{-1}\mathbf{H}_n'\mathbf{H}_n]^{-1} diag_{g=1}^{G}[\mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_{g,n})] \\
&\quad + diag_{g=1}^{G}[\boldsymbol{\alpha}_{g,n}][\boldsymbol{\Sigma}_{SQ,n}^{-1} \otimes (\mathbf{K}_n^{QQ})^{-1}] diag_{g=1}^{G}[\boldsymbol{\alpha}_{g,n}'], \\
\mathbf{S}_n^{LQ} &= \mathbf{S}_n^{LQ\prime} = \mathbf{G}_n^{QL\prime}(\boldsymbol{\Psi}_n^{QQ})^{-1}\mathbf{G}_n^{QQ} \\
&= -diag_{g=1}^{G}[\boldsymbol{\alpha}_{g,n}][\boldsymbol{\Sigma}_{SQ,n}^{-1} \otimes (\mathbf{K}_n^{QQ})^{-1}] diag_{g=1}^{G}[\mathbf{J}_{g,n}], \\
\mathbf{S}_n^{QQ} &= \mathbf{G}_n^{QQ\prime}(\boldsymbol{\Psi}_n^{QQ})^{-1}\mathbf{G}_n^{QQ} \\
&= diag_{g=1}^{G}[\mathbf{J}_{g,n}'][\boldsymbol{\Sigma}_{SQ,n}^{-1} \otimes (\mathbf{K}_n^{QQ})^{-1}] diag_{g=1}^{G}[\mathbf{J}_{g,n}].
\end{aligned}$$

Let $\widetilde{\sigma}_{gh,n}$ denote some consistent estimator for $\sigma_{gh,n}$, and let $\widetilde{\boldsymbol{\Sigma}}_n = (\widetilde{\sigma}_{gh,n})$ and $\widetilde{\boldsymbol{\Sigma}}_{SQ,n} = (\widetilde{\sigma}_{gh,n}^2)$. The submatrices of $\mathbf{S}_n$ can be estimated consistently by

$$\begin{aligned}
\widehat{\widehat{\mathbf{S}}}_n^{o,LL} &= n^{-1}\widehat{\mathbf{Z}}_{*n}'(\widehat{\widehat{\boldsymbol{\rho}}}_n^{o})(\widetilde{\boldsymbol{\Sigma}}_n^{-1} \otimes \mathbf{I}_n)\widehat{\mathbf{Z}}_{*n}(\widehat{\widehat{\boldsymbol{\rho}}}_n^{o}) \\
&\quad + diag_{g=1}^{G}[\widehat{\widehat{\boldsymbol{\alpha}}}_{g,n}^{o}][\widetilde{\boldsymbol{\Sigma}}_{SQ,n}^{-1} \otimes (\mathbf{K}_n^{QQ})^{-1}] diag_{g=1}^{G}[\widehat{\widehat{\boldsymbol{\alpha}}}_{g,n}^{o\prime}], \\
\widehat{\mathbf{S}}_{g,n}^{o,LQ} &= \widehat{\mathbf{S}}_{g,n}^{o,LQ\prime} = -diag_{g=1}^{G}[\widehat{\widehat{\boldsymbol{\alpha}}}_{g,n}^{o}][\widetilde{\boldsymbol{\Sigma}}_{SQ,n}^{-1} \otimes (\mathbf{K}_n^{QQ})^{-1}] diag_{g=1}^{G}[\mathbf{J}_{g,n}(\widehat{\widehat{\boldsymbol{\rho}}}_n^{o})], \\
\widehat{\mathbf{S}}_{g,n}^{o,QQ} &= diag_{g=1}^{G}[\mathbf{J}_{g,n}'(\widehat{\widehat{\boldsymbol{\rho}}}_n^{o})][\widetilde{\boldsymbol{\Sigma}}_{SQ,n}^{-1} \otimes (\mathbf{K}_n^{QQ})^{-1}] diag_{g=1}^{G}[\mathbf{J}_{g,n}(\widehat{\widehat{\boldsymbol{\rho}}}_n^{o})],
\end{aligned}$$

where $\widehat{\widehat{\boldsymbol{\alpha}}}_{g,n}^{o\prime} = [\widehat{\widehat{\boldsymbol{\alpha}}}_{g,1,n}^{o}, \ldots, \widehat{\widehat{\boldsymbol{\alpha}}}_{g,S,n}^{o}]'$ with

$$\widehat{\widehat{\boldsymbol{\alpha}}}_{g,r,n}^{o} = -n^{-1}\left[\mathbf{Z}_{*g,n}'(\widehat{\widehat{\boldsymbol{\rho}}}_{g,n}^{o})(\mathbf{A}_{r,n} + \mathbf{A}_{r,n}')(\mathbf{I}_n - \mathbf{R}_{g,n}^{*}(\widehat{\widehat{\boldsymbol{\rho}}}_{g,n}^{o}))\widehat{\widehat{\mathbf{u}}}_{g,n}\right]$$

and $\widehat{\widehat{\mathbf{u}}}_{g,n} = \mathbf{y}_{g,n} - \mathbf{Z}_{g,n}\widehat{\widehat{\boldsymbol{\delta}}}_n^{o}$.

F Supplemental Remarks on Identification Condition

In the following we provide a number of observations on the identification condition maintained by Assumption 6. As discussed in the text, in general, this high-level condition is a necessary condition for the identification of the regression parameters, based on linear moment conditions only. For ease of presentation we focus the discussion on moment conditions corresponding to the untransformed model (8), and we drop subscripts n. For clarity we denote the true regression parameters of the g-th equation as $\boldsymbol{\delta}_g^o$, and thus $\mathbf{u}_g = \mathbf{y}_g - \mathbf{Z}_g\boldsymbol{\delta}_g^o$.

The 2SLS estimator exploits the population moment condition $En^{-1}\mathbf{H}'\mathbf{u}_g = 0$, or asymptotically $\text{plim}_{n\to\infty} n^{-1}\mathbf{H}_n'\mathbf{u}_g = 0$. The corresponding sample moment vector satisfies

$$\begin{aligned} n^{-1}\mathbf{H}'(\mathbf{y}_g - \mathbf{Z}_g\boldsymbol{\delta}_g) &= n^{-1}\mathbf{H}'\mathbf{u}_g + n^{-1}\mathbf{H}'\mathbf{Z}_g(\boldsymbol{\delta}_g^o - \boldsymbol{\delta}_g) \\ &= \left[\text{plim}_{n\to\infty} n^{-1}\mathbf{H}'\mathbf{Z}_g\right](\boldsymbol{\delta}_g^o - \boldsymbol{\delta}_g) + o_p(1). \end{aligned}$$

Under the maintained assumptions $\mathbf{Q}_{HZ,g} = \text{plim}_{n\to\infty} n^{-1}\mathbf{H}'\mathbf{Z}_g = \lim_{n\to\infty} n^{-1}\mathbf{H}'E\mathbf{Z}_g$. Assumption 6 maintains that $\mathbf{Q}_{HZ,g}$ has full column rank. This ensures identification through the instruments $\mathbf{H}$ in that then $\boldsymbol{\delta}_g = \boldsymbol{\delta}_g^o$ is seen to be the unique solution of the $\mathbf{Q}_{HZ,g}(\boldsymbol{\delta}_g^o - \delta_g) = 0$. For ease of presentation we will proceed by discussing the finite sample analogue of the above assumption, i.e., that $\mathbf{H}'E\mathbf{Z}_g$ has full column rank. We will first relate this assumption to the rank condition in a classical simultaneous equation system, and then discuss some situations where the condition does not hold.

F.1 Interpretation from the Perspective of a Classical System

Our simultaneous equation model with network interactions contains the classical simultaneous equation model without network interactions as a special case. For the classical simultaneous equation model it is well known that a necessary and sufficient condition for the identification of the structural parameters of an equation is that the matrix of reduced form parameters corresponding to the exogenous regressors, which do not appear in that equation, has full column rank; see, e.g. Dhrymes (1978, p. 283). In the following we show that in this special case the identification condition maintained in Assumption 6 is equivalent to the classical assumption. WLOG we consider the first equation of (8), which under the classical setting reduces to

$$\mathbf{y}_1 = \mathbf{Z}_1\boldsymbol{\delta}_1 + \mathbf{u}_1,$$

where $\mathbf{Z}_1 = [\mathbf{Y}_1, \mathbf{X}_1]$ and $\boldsymbol{\delta}_1 = [\boldsymbol{\beta}_1', \boldsymbol{\gamma}_1']'$. For ease of presentation we assume furthermore that the exogenous regressors are arranged such that $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_1^*]$, where $\mathbf{X}_1^*$ represents the matrix of observation on the exogenous variables that do not appear in the first equation. Then $E\mathbf{Y}_1 = \mathbf{X}_1\Pi_1 + \mathbf{X}_1^*\Pi_1^*$, where Π_1 and Π_1^* are the $K_1 \times G_1$ and $(K - K_1) \times G_1$ matrices of reduced form parameters corresponding to $\mathbf{X}_1$ and $\mathbf{X}_1^*$ respectively. For the classical simultaneous equation model we have $\mathbf{H} = \mathbf{X}$. Thus, observing that

$$E\mathbf{Z}_1 = [E\mathbf{Y}_1, \mathbf{X}_1] = \mathbf{X}\mathbf{F}_1, \qquad \mathbf{F}_1 = \begin{bmatrix} \Pi_1 & I_{K_1} \\ \Pi_1^* & 0 \end{bmatrix},$$

we have $\mathbf{H}'E\mathbf{Z}_1 = (\mathbf{X}'\mathbf{X})\mathbf{F}_1$. Provided the $\mathbf{X}$ are not perfectly multicollinear it follows that $\mathbf{H}'E\mathbf{Z}_1$ has full column rank iff $\mathbf{F}_1$ has full column rank. However, $\mathbf{F}_1$ has full column rank iff Π_1^* has full column rank. This proves that for a classical system the assumption that $\mathbf{H}'E\mathbf{Z}_1$ is equivalent to the classical rank condition. Of course, the order condition $K - K_1 \geq G_1$ is necessary for the rank condition to hold.

F.2 Interpretations from a Stylized Two-Equation Model with Network Interactions

In the following we consider the simple stylized two-equation model:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{y}_2\beta_1 + \mathbf{X}_1\boldsymbol{\gamma}_1 + \lambda_1\mathbf{W}\mathbf{y}_1 + \mathbf{u}_1 = \mathbf{Z}_1\boldsymbol{\delta}_1 + \mathbf{u}_1, \\ \mathbf{y}_2 &= \mathbf{X}_2\boldsymbol{\gamma}_2 + \mathbf{u}_2, \end{aligned}$$

where $\mathbf{Z}_1 = [\mathbf{y}_2, \mathbf{X}_1, \mathbf{W}\mathbf{y}_1]$ and $\boldsymbol{\delta}_1 = [\beta_1, \boldsymbol{\gamma}_1', \lambda_1]'$. While the setup is simple it permits us to highlight several important scenarios where the identification condition for the parameters of the first equation fails.

We first compute the best instruments $E\mathbf{Z}_1$ for $\mathbf{Z}_1$. Observe that

$$\begin{aligned} \mathbf{y}_1 &= (\mathbf{X_2}\boldsymbol{\gamma_2} + \mathbf{u_2})\beta_1 + \mathbf{X}_1\boldsymbol{\gamma}_1 + \lambda_1\mathbf{W}\mathbf{y}_1 + \mathbf{u}_1 \\ &= \mathbf{X}\pi_1 + \lambda_1\mathbf{W}\mathbf{y}_1 + \mathbf{u}_1 + \mathbf{u_2}\beta_1 \end{aligned}$$

with $\boldsymbol{\pi}_1 = (\boldsymbol{\gamma}_1', \beta_1\boldsymbol{\gamma}_2')'$, and the reduced form for $\mathbf{y}_1$ is given by

$$\mathbf{y}_1 = (\mathbf{I}_n - \lambda_1\mathbf{W})^{-1}[\mathbf{X}\pi_1 + \mathbf{u}_1 + \mathbf{u}_2\beta_1].$$

Consequently the best instruments $E\mathbf{Z}_1$ for $\mathbf{Z}_1$ is given by

$$\begin{aligned} E\mathbf{Z}_1 = [E\mathbf{y}_2, \mathbf{X}_1, \mathbf{W}E\mathbf{y}_1] &= [\mathbf{X}_2\boldsymbol{\gamma}_2, \mathbf{X}_1, \mathbf{W}(\mathbf{I}_n - \lambda_1\mathbf{W})^{-1}\mathbf{X}\pi_1] \\ &= [\mathbf{X}_2\boldsymbol{\gamma}_2, \mathbf{X}_1, \sum_{s=0}^{\infty}\lambda_1^s\mathbf{W}^{s+1}\mathbf{X}\pi_1]. \end{aligned}$$

The best IV estimator for $\boldsymbol{\delta}_1$ is then given by

$$\widehat{\boldsymbol{\delta}}_1^B = [(E\mathbf{Z}_1)'\mathbf{Z}_1]^{-1}(E\mathbf{Z}_1)'\mathbf{y}_1.$$

From the above we see that the moment condition $E(E\mathbf{Z}_1)'\mathbf{u}_1 = 0$ corresponding to the best instruments is a weighted average of the basic underlying moment conditions

$$E(\mathbf{W}^s\mathbf{X})'\mathbf{u}_1 = 0, \ldots, s = 0, 1, \ldots, \infty. \tag{F.1}$$

The best instruments $E\mathbf{Z}_1$ depend on the inverse of the $n \times n$ matrix $\mathbf{I}_n - \lambda_1\mathbf{W}$, which may be computationally challenging in large samples. In light of (F.1), adapting Kelejian and Prucha (1998), we can define $\mathbf{H}$ to be composed of the linearly independent columns of $\mathbf{W}\mathbf{X}, \ldots, \mathbf{W}^S\mathbf{X}$ for some $S \geq 1$, and work with the moment condition

$$E\mathbf{H}'\mathbf{u}_1 = 0.$$

The corresponding optimal GMM estimator is the S2SLS estimator

$$\widehat{\boldsymbol{\delta}}_1 = [(\widehat{\mathbf{Z}}_1)'\mathbf{Z}_1]^{-1}(\widehat{\mathbf{Z}}_1)'\mathbf{y}_1$$

with $\widehat{\mathbf{Z}}_1 = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\mathbf{Z}_1$, and where $\widehat{\mathbf{Z}}_1$ can be viewed to represent an approximation of the ideal instruments.

Kelejian and Prucha (1998) discuss identification based on linear moment restrictions for a single equation cross sectional spatial model. In line with their discussion, clearly for $\mathbf{H}'E\mathbf{Z}_1$ to have full column rank we need $E\mathbf{Z}_1$ to have full column rank. To provide guidance on where this condition may fail, we next discuss several such scenarios.

■ Scenario 1: Suppose the true model does not contain any exogenous variables, i.e., $\boldsymbol{\gamma}_1 = 0$ and $\gamma_2 = 0$. Obviously in this case $\boldsymbol{\pi}_1 = 0$ and $\mathbf{WEy}_1 = 0$, and thus $E\mathbf{Z}_1$ does not have full column rank. Apart from a complete failure of identification by linear moments under this scenario, we expect the instruments $\mathbf{X}$, $\mathbf{WX}$, $\mathbf{W}^2\mathbf{X}$, $\ldots$ to be weak, and estimators based only on linear moment conditions to perform poorly, when the parameters of the exogenous variables are "small". Since the values of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ depend on the chosen units of measurement of the exogenous variables, it seems intuitive that "small" is best interpreted as to correspond to a small ratio of the variance/signal stemming from the exogenous variables to the variance/noise of the disturbances.

■ Scenario 2: Suppose $\mathbf{W}$ is such that $\mathbf{W}^2 = c_1\mathbf{I}_n + c_2\mathbf{W}$, which implies that $(\mathbf{I}_n - \lambda_1\mathbf{W})^{-1} = d_1\mathbf{I}_n + d_2\mathbf{W}$ where the constants d_1 and d_2 depend on λ_1, c_1, c_2.[1] Suppose further that $\mathbf{X}_1 = [\underline{\mathbf{X}}_1, \mathbf{W}\underline{\mathbf{X}}_1]$, and correspondingly $\boldsymbol{\gamma}_1 = [\boldsymbol{\gamma}_1^{x\prime}, \boldsymbol{\gamma}_1^{w\prime}]'$, and that $\beta_1 = 0$. In this case we have

$$\begin{aligned}
\mathbf{W}E\mathbf{y}_1 &= \mathbf{W}(\mathbf{I}_n - \lambda_1\mathbf{W})^{-1}\mathbf{X}\pi_1 = (d_1\mathbf{W} + d_2\mathbf{W}^2)\mathbf{X}\pi_1 \\
&= [c_1 d_2\mathbf{I}_n + (d_1 + d_2 c_2)\mathbf{W}][\underline{\mathbf{X}}_\mathbf{1}\boldsymbol{\gamma}_\mathbf{1}^\mathbf{x} + \mathbf{W}\underline{\mathbf{X}}_\mathbf{1}\gamma_1^w] \\
&= c_1 d_2\underline{\mathbf{X}}_\mathbf{1}\boldsymbol{\gamma}_1^x + (d_1 + d_2 c_2)\mathbf{W}\underline{\mathbf{X}}_\mathbf{1}\gamma_1^x \\
&\quad + c_1 d_2\mathbf{W}\underline{\mathbf{X}}_\mathbf{1}\boldsymbol{\gamma}_1^w + (d_1 + d_2 c_2)[c_1\mathbf{I}_n + c_2\mathbf{W}]\underline{\mathbf{X}}_\mathbf{1}\gamma_1^w \\
&= c_1 d_2\underline{\mathbf{X}}_\mathbf{1}\boldsymbol{\gamma}_1^x + (d_1 + d_2 c_2)c_1\underline{\mathbf{X}}_\mathbf{1}\boldsymbol{\gamma}_1^w \\
&\quad + (d_1 + d_2 c_2)\mathbf{W}\underline{\mathbf{X}}_\mathbf{1}\gamma_1^x + c_1 d_2\mathbf{W}\underline{\mathbf{X}}_\mathbf{1}\boldsymbol{\gamma}_1^w + (d_1 + d_2 c_2)c_2\mathbf{W}\underline{\mathbf{X}}_\mathbf{1}\gamma_1^w
\end{aligned}$$

which is clearly collinear with the columns of $\mathbf{X}_1$. We note that the result is specific. It would generally not extend to the case where $\mathbf{X}_1$ only contains a subset of spatial lags of the exogenous variables. It would generally also not extend to the case where $\beta_1 \neq 0$, i.e., to the case where additional endogenous variables are present that would depend on additional exogenous variables that can serve as instruments.

A leading example where $\mathbf{W}^2 = c_1\mathbf{I}_n + c_2\mathbf{W}$, and instrumentation in terms of neighbor's characteristics fails arises if there are R groups of size m_g, $g = 1, \ldots, R$, and social interactions take place only within groups, and all members of a group are friends of equal importance. If the calculation of group means includes all members we have $\mathbf{W} = diag_{g=1}^R(\mathbf{W}_{m_g})$ with $\mathbf{W}_{m_g} = e_{m_g}e_{m_g}'/m_g$, where e_{m_g} denotes an $m_g \times 1$ vector of ones. If the calculation of group means affecting the i-th member excludes the i-th member we have $\mathbf{W} = diag_{g=1}^R(\mathbf{W}_{m_g})$

[1] This is readily verified by observing that $\mathbf{I}_n = (\mathbf{I}_n - \lambda_1\mathbf{W})(d_1\mathbf{I}_n + d_2\mathbf{W})$ and utilizing the expression for $\mathbf{W}^2$.

with $\mathbf{W}_{m_g} = (e_{m_g}e'_{m_g} - I_{m_g})/(m_g - 1)$. Both in the first case and, provided that all groups are of the same size, also in the second case we have $\mathbf{W}^2 = c_1\mathbf{I}_n + c_2\mathbf{W}$ and identification via instruments fails. However, in the latter case identification is achievable if there is variation in the group size. For a further discussion of these cases for cross sectional data see Bramoulle, Djebbari and Fortin (2009) and Paula (2017), and Kelejian and Prucha (2002) and Kelejian et al. (2006) for an early discussion of identification in case of equal weights.

■ Scenario 3: Consider the model

$$\mathbf{y}_1 = \underline{\mathbf{X}}_{\mathbf{1}}\boldsymbol{\gamma}_1^x + \mathbf{W}\underline{\mathbf{X}}_1\boldsymbol{\gamma}_1^w + \lambda_1\mathbf{W}\mathbf{y}_1 + \mathbf{u}_1 - \lambda_1\mathbf{W}\mathbf{u}_1.$$

Then $E\mathbf{Z}_1 = [\underline{\mathbf{X}}_{\mathbf{1}}, \boldsymbol{W}\underline{\mathbf{X}}_1, \mathbf{W}E\mathbf{y}_1]$ does not have full column rank for parameter constellations where $\boldsymbol{\gamma}_1^w = -\lambda_1\boldsymbol{\gamma}_1^x$. To see this observe that for those parameter constellations $E\mathbf{y}_1 = \underline{\mathbf{X}}_{\mathbf{1}}\boldsymbol{\gamma}_1^x$. For interpretation, we note that the above model is observationally equivalent to $\mathbf{y}_1 = \underline{\mathbf{X}}_{\mathbf{1}}\boldsymbol{\gamma}_1^x + \mathbf{u}_1$, since pre-multiplication of this model with $\mathbf{I}_n - \lambda_1\mathbf{W}$ yields the above model.

F.3 Identification from Linear and Quadratic Moment Conditions

A standard assumption in the literature on GMM estimation is that the probability limit of the matrix of first order derivatives of the moment vector w.r.t. to the parameters has full column rank when evaluated at the true parameter (or a small sample analogue thereof). Let $\mathbf{m}_g^{\delta,u}(\boldsymbol{\delta}_g) = n^{-1}\mathbf{H}'\mathbf{u}_g(\boldsymbol{\delta}_g)$ with $\mathbf{u}_g(\boldsymbol{\delta}_g) = \mathbf{y}_g - \mathbf{Z}_g\boldsymbol{\delta}_g$. Then $\text{plim}_{n\to\infty}\partial\mathbf{m}_g^{\delta,u}(\boldsymbol{\delta}_g^o)/\partial\boldsymbol{\delta}_g = \mathbf{Q}_{HZ,g}$. Furthermore, let $\mathbf{m}_g^{\delta}(\boldsymbol{\rho}_g, \boldsymbol{\delta}_g)$ and $\mathbf{m}_g^{\rho}(\boldsymbol{\rho}_g, \boldsymbol{\delta}_g)$ be defined as in (12) and recall the equivalent definition of the latter in (14). Then, $\text{plim}_{n\to\infty}\partial\mathbf{m}_g^{\delta}(\boldsymbol{\rho}_g^o, \boldsymbol{\delta}_g^o)/\partial\boldsymbol{\delta}_g = \mathbf{Q}_{HZ,g*}(\boldsymbol{\rho}_g)$ and $\text{plim}_{n\to\infty}\partial\mathbf{m}_g^{\rho}(\boldsymbol{\rho}_g, \boldsymbol{\delta}_g)/\partial\boldsymbol{\rho}_g = \boldsymbol{\Gamma}_g[\partial\mathbf{r}_g/\partial\boldsymbol{\rho}_g]$. From this we see that the nature of Assumptions 6 and 7 is in line with assumptions maintained by the classical GMM literature; observe that $\partial\mathbf{r}_g/\partial\boldsymbol{\rho}_g$ has full column rank.

As remarked in the text, the conditions postulated in Assumption 6 are sufficient conditions to ensure identification of the regression parameters $\boldsymbol{\delta}_g$ from the linear moments only. Given $\boldsymbol{\delta}_g$ is identified, the conditions postulated in Assumption 7 are sufficient to ensure the identification of the autoregressive parameters $\boldsymbol{\rho}_g$. Assumptions 6 and 7 are geared towards two-step estimation.

Within the context of one-step estimation identification is still possible with the use of the quadratic moment conditions, even if identification by the linear moment conditions fails. In this case a sufficient condition for the parameters $\boldsymbol{\theta}_g = [\boldsymbol{\rho}_g', \boldsymbol{\delta}_g']'$ to be identified is that $\text{plim}_{n\to\infty}\partial\mathbf{m}_g(\boldsymbol{\theta}_g^o)/\partial\boldsymbol{\theta}_g$ has full column rank, where $\mathbf{m}_g(\boldsymbol{\theta}_g)$ denotes the stacked vector of linear and quadratic moment conditions. For contributions on identification with the help of quadratic moment conditions that ensure that $\text{plim}_{n\to\infty}\partial\mathbf{m}_g(\boldsymbol{\theta}_g^o)/\partial\boldsymbol{\theta}_g$ has full column rank see, e.g., Lee (2007a) and Kuersteiner and Prucha (2020) within a single equation framework, and see, e.g., Liu (2014, 2019, 2020), Liu and Saraiva (2019), and Yang and Lee (2017,2019) for contributions within a systems framework. These contributions focus on the case where the disturbances are uncorrelated in the cross section, i.e. where $\boldsymbol{\theta}_g = \boldsymbol{\delta}_g$.

G Monte Carlo Study

To analyze the small sample properties of our estimators we have conducted a Monte Carlo study. The study considers both weights matrices motivated by social interactions and spatial interactions. The study explores situations where the parameters are identified by the linear moment conditions alone. It also considers situations where the parameters are identified from utilizing both linear and quadratic moment conditions jointly, but where identification from the linear moment conditions alone is weak.

In Appendix G.1 we describe the Monte Carlo design and provide highlights of the study for two-step estimators for the case where the parameters are identified by linear moment conditions. The simulation results in Appendix G.1 are based on weights matrices corresponding to an underlying social interactions structure. In Appendices G.2 and G.3 we report on additional Monte Carlo simulations for the identified case for weights matrices motivated by social interactions as well as by spatial interactions. We also report on results for the three scenarios where identification is weak discussed in the Appendix F included in this Online Supplementary Appendix.

G.1 Monte Carlo Design and Main Results

For the Monte Carlo results below we considered the following two equation system as a special case of (2):

$$\begin{aligned}
&\mathbf{y}_1 = b_{21}\mathbf{y}_2 + [\lambda_{11,1}\mathbf{M}_1 + \lambda_{11,2}\mathbf{M}_2]\mathbf{y}_1 + \textstyle\sum_{k=1}^{3} c_{k1}\mathbf{x}_k + \mathbf{u}_1,\\
&\mathbf{y}_2 = b_{12}\mathbf{y}_1 + [\lambda_{22,1}\mathbf{M}_1 + \lambda_{22,2}\mathbf{M}_2]\mathbf{y}_2 + \textstyle\sum_{k=4}^{6} c_{k2}\mathbf{x}_k + \mathbf{u}_2,\\
&\mathbf{u}_g = [\rho_{g1}\mathbf{M}_1 + \rho_{g2}\mathbf{M}_2]\mathbf{u}_g + \boldsymbol{\epsilon}_g, \qquad g = 1, 2.
\end{aligned} \tag{G.1}$$

The stylized social-network design employes a group structure, which can be viewed as emulating groups of friends in a classroom setting. More specifically, suppose there are P schools, and each school has three classrooms of size m_1, m_2, and m_3. Now consider the matrix

$$E = (e_{ij}) = I_P \otimes \begin{bmatrix} E_{m_1} & 0 & 0 \\ 0 & E_{m_2} & 0 \\ 0 & 0 & E_{m_3} \end{bmatrix}$$

where the E_{m_q}, $q = 1, 2, 3$, are $m_q \times m_q$ matrices with zeros on the diagonal and with ones off the diagonal. Then the elements of E can be viewed as indicator variables that are equal to one if two students belong to the same classroom, and zero otherwise. The matrix E is of dimension $n \times n$, where the sample size is $n = mP$ with $m = m_1 + m_2 + m_3$. We used the following values in our Monte Carlo simulations:

$$m_1 = 10,\ m_2 = 15,\ m_3 = 25,\ P = 10 \text{ or } 20,$$

which implies a sample size of $n = 500$ or $1,000$.

To generate the weights matrices $\mathbf{M}_1$ and $\mathbf{M}_2$, let ξ_{Gi} and ξ_{Ii} be, respectively, an i.i.d. binary random variable taking values 0 and 1 with equal probability, and an i.i.d. discrete random variable taking values $1, 2, \ldots, 10$ with equal probability, and let μ_i be i.i.d. $N(0,1)$. Furthermore, (ξ_{Gi}), (ξ_{Ii}) and (μ_i) are generated independently. Now define

$$d_{ij} = \left[c_G(\xi_{Gi} - \xi_{Gj})/\sigma_G + c_I(\xi_{Ii} - \xi_{Ij})/\sigma_I + c_\mu(\mu_i - \mu_j)\right]/\sigma_\mu$$

where $\sigma_G^2 = 1/4$, $\sigma_I^2 = (100-1)/12$, $\sigma_\mu^2 = 1$ denote the variances of ξ_{Gi}, ξ_{Ii}, and μ_i, respectively. For an exemplary interpretation, ξ_{Gi} could be an indicator for the gender of an individual, ξ_{Ii} could represent the family income decile of an individual, the μ_i are unobserved characteristics, and d_{ij} could then be interpreted as a measure of similarity between two individuals. With this interpretation we define

$$\begin{aligned} m_{1,ij}^* &= \mathbf{1}(|d_{ij}| < d_*)e_{ij}, \\ m_{2,ij}^* &= \mathbf{1}(d_* \le |d_{ij}| < d_{**})e_{ij}, \end{aligned}$$

where $m_{1,ij}^*$ and $m_{2,ij}^*$ are now indicators that equal one, respectively, if two individuals are best friends or just friends, and zero otherwise.

Specific values that generate on average about 29% best friends and 40% friends are:

$$c_G = .4,\ c_I = .4,\ c_\mu = .2,\ d_* = .3,\ d_{**} = .8.$$

The weights matrices $\mathbf{M}_1 = [m_{1,ij}]$ and $\mathbf{M}_2 = [m_{2,ij}]$ are then obtained by applying the following normalization ($s = 1, 2$): $m_{s,ij} = m_{s,ij}^* / \sum_{j=1}^n m_{s,ij}^*$ if $\sum_{j=1}^n m_{s,ij}^* > 0$ and $m_{s,ij} = m_{s,ij}^*$ if $\sum_{j=1}^n m_{s,ij}^* = 0$. Note that the design allows for situations where an individual has no close friends or only close friends. That is, we allow for a row of $\mathbf{M}_1$ or $\mathbf{M}_2$ to only contain zeros. If a row contains nonzero elements, then that row is normalized so that the row sum is one.

We consider three sets of parameters of model (G.1). Set I corresponds to positive spillovers, Set II to negative spillovers and Set III to zero spillovers. In particular, we consider the parameter values as given in Table 1:

Table 1: Configuration of Autoregressive Parameters in Set I-III

	Auroregressive Parameters							
	Equation 1				Equation 2			
	$\lambda_{11,1}$	$\lambda_{11,2}$	ρ_{11}	ρ_{12}	$\lambda_{22,1}$	$\lambda_{22,2}$	ρ_{21}	ρ_{22}
Set I	0.30	0.20	0.20	0.10	0.30	0.15	0.10	0
Set II	-0.30	-0.20	-0.20	-0.10	-0.30	-0.15	-0.10	0
Set III	0	0	0	0	0	0	0	0

The remaining parameters are selected as $b_{12} = 0.3$, $b_{21} = 0.15$, $c_{1k} = 1$ for $k = 1, 2, 3$, and $c_{2k} = 1$ for $k = 4, 5, 6$. The observations on the exogenous regressors $\mathbf{x}_1, \ldots, \mathbf{x}_6$ are kept fixed for all Monte Carlo iterations, and are generated as independent of each other and as cross sectionally *i.i.d.* N(1,3). The disturbances $\boldsymbol{\epsilon}_1$, $\boldsymbol{\epsilon}_2$ are generated as cross sectionally i.i.d. normal with mean 0, variance 1 and covariance .5.

In Table 2, given at the end of this subsection, we report on the bias and root mean squared error (RMSE) of the maximum-likelihood estimator (ML), the GS2SLS, and the GS3SLS estimator for parameter Set I based on $1,000$ Monte Carlo repetitions.[2] More specifically, to simplify the presentation, in Table 2 we only report on a subset of the parameters of the first equation of Model (G.1), corresponding to $\mathbf{y}_2$ as well as on the autoregressive parameters. For all estimators the biases are fairly small, indicating that the linear moments alone are able to identify the regression parameters. As expected, the ML estimator has the smallest RMSE. In general, in terms of RMSE, the ML only dominates GS3SLS slightly, and GS3SLS dominates GS2SLS. The differences in RMSE are the most pronounced for the estimates of the autoregressive parameters in the

[2] Our measure of bias is defined as the difference between the median and the true parameter value. Our measure corresponding to the RMSE is defined as $[bias^2 + (IQ/1.35)^2]^{1/2}$ where IQ is the inter-quantile range. That is, $IQ = c_1 - c_2$ where c_1 is the .75 quantile and c_2 is the .25 quantile. If the distribution is normal, $IQ/1.35$ is (apart from rounding errors) equal to the standard deviation. In the following we will refer to our measures simply as bias and RMSE.

disturbance process. As expected, biases and RMSE decline with the sample size.

The results in Table 2 represent a subset of the results reported in Tables 4-6 in Appendix G.2. In Tables 4-6 we report on the estimators of all the parameters of both equations for parameter Sets I-III. In addition to considering the above described scenarios with social network interactions, we also report in Tables 7-9 in Appendix G.2 on results from scenarios with spatial network interactions, using a spatial rook design. In general, the results are in line with those reported above. In the Appendix G.3 we also cover several scenarios where identification by linear moment conditions alone is weak, and where in consequence GS2SLS and GS3SLS can be substantially biased. Our extended Monte Carlo results also report on the performance of LQ-GS2SLS and LQ-GS3SLS. Under weak identification LQ-GS2SLS and LQ-GS3SLS can greatly outperform GS2SLS and GS3SLS. However, for the well identified scenarios underlying the results in Tables 4-9 the benefit of combining linear and quadratic moment conditions seems limited.

A leading hypothesis of interest is the absence of spillovers. Focusing on equation 1 we can test the hypothesis $H_0 : \lambda^0_{11,1} = \lambda^0_{11,2} = \rho^0_{11} = \rho^0_{12} = 0$. In Table 3, given at the end of this subsection, we report on the power of the Wald test for this hypothesis based on GS2SLS and GS3SLS estimates. More specifically, we explore the power of the test for parameter values $(\lambda^0_{11,1}, \lambda^0_{11,2}, \rho^0_{11}, \rho^0_{12}) = \kappa(\lambda^I_{11,1}, \lambda^I_{11,2}, \rho^I_{11}, \rho^I_{12})$ where $(\lambda^I_{11,1}, \lambda^I_{11,2}, \rho^I_{11}, \rho^I_{12}) = (.30, .30, .20, .10)$ is equal to the values considered by parameter Set I, and where the factor $\kappa = 0, 0.05, .., 0.30$. All other parameter values are kept as in parameter Set I. The significance level of the tests are close to the nominal 5 percent, especially for sample size $n = 1,000$, and the tests seem to have good power.

Table 2: Bias and RMSE of MLE, GS2SLS and GS3SLS Parameters of Equation 1, Social Interaction Weights Matrices

		Soc. Interact. Matrices; $n = 500$						Soc. Interact. Matrices; $n = 1000$					
		MLE		GS2SLS		GS3SLS		MLE		GS2SLS		GS3SLS	
Parameter	True	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Eq. 1	Set I												
b_{21}	0.15	0.00047	0.01372	0.00304	0.01358	0.00074	0.01357	-0.00004	0.00973	0.00175	0.01011	0.00031	0.00978
$\lambda_{11,1}$	0.3	0.00030	0.01556	0.00156	0.01708	0.00281	0.01551	-0.00089	0.01129	-0.00084	0.01276	0.00016	0.01147
$\lambda_{11,2}$	0.2	-0.00028	0.01651	-0.00366	0.01849	-0.00209	0.01800	0.00072	0.01093	-0.00029	0.01202	-0.00007	0.01121
ρ_{11}	0.2	-0.00536	0.05467	-0.00261	0.05724	-0.00271	0.05807	-0.00157	0.03793	0.00074	0.04028	0.00021	0.04068
ρ_{12}	0.1	-0.00113	0.06185	0.00162	0.07189	0.00137	0.06941	-0.00338	0.04305	-0.00043	0.04806	-0.00046	0.04829
Eq. 1	Set II												
b_{21}	0.15	0.00051	0.01319	0.00451	0.01360	0.00208	0.01333	-0.00005	0.00963	0.00225	0.01002	0.00095	0.00952
$\lambda_{11,1}$	-0.3	0.00057	0.02187	-0.00054	0.02364	-0.00017	0.02262	-0.00124	0.01310	-0.00185	0.01475	-0.00152	0.01313
$\lambda_{11,2}$	-0.2	0.00013	0.02256	-0.00204	0.02572	-0.00096	0.02267	0.00160	0.01311	0.00039	0.01462	0.00053	0.01359
ρ_{11}	-0.2	-0.00583	0.06664	0.00098	0.06941	-0.00040	0.07170	-0.00098	0.04289	0.00198	0.04644	0.00138	0.04717
ρ_{12}	-0.1	-0.00048	0.07936	0.00398	0.08989	0.00403	0.09009	-0.00279	0.05461	-0.00069	0.06176	-0.00095	0.05952
Eq. 1	Set III												
b_{21}	0.15	0.00051	0.01351	0.00411	0.01367	0.00153	0.01336	-0.00019	0.00980	0.00216	0.01014	0.00073	0.00982
$\lambda_{11,1}$	0	0.00064	0.02031	0.00157	0.02164	0.00258	0.01980	-0.00137	0.01282	-0.00125	0.01434	-0.00053	0.01274
$\lambda_{11,2}$	0	-0.00035	0.02179	-0.00331	0.02369	-0.00211	0.02164	0.00117	0.01311	0.00020	0.01429	0.00034	0.01296
ρ_{11}	0	-0.00538	0.06453	-0.00145	0.06674	-0.00351	0.06660	-0.00084	0.04160	0.00055	0.04401	-0.00010	0.04533
ρ_{12}	0	-0.00066	0.07220	0.00343	0.08182	0.00414	0.08272	-0.00404	0.05041	-0.00108	0.05518	-0.00073	0.05425

Table 3: Power Function of Joint Wald Tests Corresponding to GS2SLS and GS3SLS for $H_0 : \lambda^0_{11,1} = \lambda^0_{11,2} = \rho^0_{11} = \rho^0_2 = 0$. The True Autoregressive Parameters Equal Those of Set I Scaled by κ. The Significance Level is Displayed at $\kappa = 0$.

	Social Interactions Matrices			
	$n = 500$		$n = 1000$	
κ	GS2SLS	GS3SLS	GS2SLS	GS3SLS
0	0.068	0.072	0.058	0.054
0.05	0.129	0.170	0.226	0.272
0.1	0.423	0.514	0.773	0.840
0.15	0.832	0.889	0.993	0.999
0.2	0.987	0.994	1	1
0.25	1	1	1	1
0.3	1	1	1	1

G.2 Additional Monte Carlo Simulations for the Identified Case

Tables 4-6, given at the end of this subsection, are an expansion of Table 2. In Tables 4-6 we report additionally on the performance of the one-step LQ-GS2SLS and LQ-GS3SLS estimators, as well as on a larger set of parameters. More specifically, in Tables 4-6 we provide information on the bias and RMSE for various estimators of the parameters in equations 1 and 2 of Model (G.1) for parameter Sets I-III and sample sizes $n = 500$ and $n = 1,000$ each. Tables 4-6 are based on the same design of the social network weights matrices as those underlying Table 2.

The results in Tables 4-6 are in line with the subset of results reported in Table 2. For all estimators the biases are fairly small, indicating that the linear moments alone are able to identify the regression parameters. As expected, the ML estimator has the smallest RMSE. In general, in terms of RMSE, the ML only dominates GS3SLS slightly, and GS3SLS dominates G2SLS. The differences in RMSE are the most pronounced for the estimates of the autoregressive parameters in the disturbance process. The differences in RMSE for the parameters of the exogenous parameters are especially small. As expected, biases and RMSE decline with the sample size. In Tables 4-6 we also report on the performance of the LQ-GS2SLS and LQ-GS3SLS. While LQ-GS2SLS and LQ-GS3SLS have the potential to greatly outperform GS2SLS and GS3SLS under weak identification, for the well identified scenarios underlying the results in Tables 4-6 the benefit of combining linear and quadratic moment conditions seems limited.

In Tables 7-9, given at the end of this subsection, we provide information on the bias and RMSE for the same estimators for the parameters of Model (G.1) as in Tables 4-6, but for an alternative set of weights matrices. We explore the same parameter specifications, but now generate the data based on weights matrices "inspired" by a spatial network. More specifically, we derive the weights matrices from a classical rook design. In more detail, to define locations and neighbors, consider a square grid with both the x and y coordinates only taking on the values $1, 2, ..., m$. Next, define the Euclidean distance between any pair of units, i_1 and i_2 with coordinates (x_1, y_1) and (x_2, y_2), respectively, as $d(i_1, i_2) = \left[(x_1 - x_2)^2 + (y_1 - y_2)^2\right]^{1/2}$. Moreover, define the cutoff distances $d_* = 1$ and $d_{**} = 2$ to determine inner- and outer-ring neighbors, respectively, around any spatial unit on the lattice. Now define the (i, j)-th element of a row-normalized weights matrix $\mathbf{M}_1$ and $\mathbf{M}_2$ as

$$m_{r,ij} = m^*_{r,ij} / \sum_{j=1}^{n} m^*_{r,ij}, \qquad r \in 1, 2, \tag{G.2}$$

$$m^*_{1,ij} = \begin{cases} 1 & \text{if} \quad 0 < d(i_1, i_2) \leq d_* \\ 0 & \text{else} \end{cases}, \tag{G.3}$$

$$m^*_{2,ij} = \begin{cases} 1 & \text{if} \quad d_* < d(i_1, i_2) \leq d_{**} \\ 0 & \text{else} \end{cases}. \tag{G.4}$$

We consider two configurations. The two configurations correspond to $m = 22$ and $m = 31$, which implies a sample size of $n = 484$ and $n = 961$, respectively. The finding in Tables 7-9 are similar to those in Tables 4-6.

Table 4: Bias and RMSE of Various Estimators for Parameters of Equation 1 and 2, Parameter Set I with Social Interaction Weights Matrices

		Soc. Interact. Weights Matrices; $n = 500$									
Parameter	True	MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	Set I	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	0.00047	0.01372	0.00304	0.01358	0.00074	0.01357	0.00538	0.01459	0.00256	0.01356
$\lambda_{11,1}$	0.3	0.00030	0.01556	0.00156	0.01708	0.00281	0.01551	-0.00178	0.01656	-0.00076	0.01570
$\lambda_{11,2}$	0.2	-0.00028	0.01651	-0.00366	0.01849	-0.00209	0.01800	-0.00294	0.01911	-0.00146	0.01784
ρ_{11}	0.2	-0.00536	0.05467	-0.00261	0.05724	-0.00271	0.05807	0.03205	0.07140	0.03635	0.07269
ρ_{12}	0.1	-0.00113	0.06185	0.00162	0.07189	0.00137	0.06941	0.01227	0.07803	0.01232	0.07579
c_{11}	1	-0.00072	0.02415	-0.00124	0.02677	-0.00131	0.02407	-0.00090	0.02739	-0.00065	0.02393
c_{12}	1	-0.00089	0.02580	-0.00385	0.02596	-0.00189	0.02559	-0.00289	0.02587	-0.00218	0.02561
c_{13}	1	0.00034	0.02428	-0.00288	0.02531	-0.00040	0.02425	-0.00336	0.02558	-0.00037	0.02454
Eq. 2	Set I										
b_{12}	0.3	-0.00068	0.01335	0.00219	0.01356	-0.00121	0.01331	0.00388	0.01374	0.00031	0.01320
$\lambda_{22,1}$	0.3	0.00061	0.01451	0.00194	0.01540	0.00271	0.01496	-0.00041	0.01579	0.00030	0.01509
$\lambda_{22,2}$	0.15	-0.00052	0.01457	-0.00321	0.01693	-0.00130	0.01547	-0.00199	0.01685	-0.00093	0.01584
ρ_{21}	0.1	-0.01171	0.05897	-0.00924	0.06391	-0.00992	0.06379	0.01850	0.07065	0.02952	0.07403
ρ_{22}	0	-0.00412	0.06792	-0.00091	0.07142	-0.00213	0.07009	0.00954	0.07986	0.01287	0.07974
c_{24}	1	-0.00118	0.02173	-0.00160	0.02191	-0.00157	0.02139	-0.00112	0.02288	-0.00011	0.02193
c_{25}	1	0.00175	0.02423	-0.00100	0.02605	0.00190	0.02437	-0.00068	0.02658	0.00222	0.02438
c_{26}	1	0.00192	0.02268	0.00005	0.02285	0.00128	0.02266	0.00008	0.02326	0.00207	0.02313
		$n = 1000$									
Eq. 1	Set I										
b_{21}	0.15	-0.00004	0.00973	0.00175	0.01011	0.00031	0.00978	0.00264	0.01025	0.00138	0.00968
$\lambda_{11,1}$	0.3	-0.00089	0.01129	-0.00084	0.01276	0.00016	0.01147	-0.00221	0.01308	-0.00132	0.01161
$\lambda_{11,2}$	0.2	0.00072	0.01093	-0.00029	0.01202	-0.00007	0.01121	0.00038	0.01202	0.00039	0.01140
ρ_{11}	0.2	-0.00157	0.03793	0.00074	0.04028	0.00021	0.04068	0.01550	0.04566	0.01838	0.04524
ρ_{12}	0.1	-0.00338	0.04305	-0.00043	0.04806	-0.00046	0.04829	0.00305	0.04854	0.00448	0.04896
c_{11}	1	-0.00001	0.01585	-0.00108	0.01704	-0.00060	0.01586	-0.00132	0.01711	-0.00011	0.01604
c_{12}	1	-0.00026	0.01632	-0.00130	0.01798	-0.00079	0.01634	-0.00146	0.01786	-0.00016	0.01630
c_{13}	1	-0.00025	0.01661	-0.00107	0.01858	-0.00050	0.01661	-0.00136	0.01911	-0.00033	0.01636
Eq. 2	Set I										
b_{12}	0.3	-0.00001	0.00912	0.00140	0.00949	0.00003	0.00941	0.00198	0.00953	0.00080	0.00931
$\lambda_{22,1}$	0.3	0.00034	0.00957	0.00106	0.01111	0.00152	0.01020	-0.00025	0.01136	-0.00002	0.01017
$\lambda_{22,2}$	0.15	0.00012	0.01052	-0.00231	0.01154	-0.00107	0.01058	-0.00158	0.01154	-0.00041	0.01074
ρ_{21}	0.1	-0.00593	0.04269	-0.00494	0.04359	-0.00404	0.04287	0.00970	0.04726	0.01557	0.04763
ρ_{22}	0	-0.00424	0.04485	-0.00132	0.04997	-0.00365	0.04912	-0.00089	0.05335	0.00104	0.05281
c_{24}	1	0.00051	0.01672	-0.00047	0.01858	0.00056	0.01685	-0.00075	0.01876	0.00089	0.01671
c_{25}	1	0.00004	0.01698	-0.00065	0.01838	0.00003	0.01667	-0.00056	0.01795	0.00019	0.01634
c_{26}	1	-0.00046	0.01745	-0.00087	0.01894	-0.00037	0.01749	-0.00067	0.01843	-0.00005	0.01771

Table 5: Bias and RMSE of Various Estimators for Parameters of Equation 1 and 2, Parameter Set II with Social Interaction Weights Matrices

		Soc. Interact. Weights Matrices; $n = 500$									
Parameter	True	MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	Set II	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	0.00051	0.01319	0.00451	0.01360	0.00208	0.01333	0.00522	0.01390	0.00320	0.01369
$\lambda_{11,1}$	-0.3	0.00057	0.02187	-0.00054	0.02364	-0.00017	0.02262	-0.00193	0.02467	-0.00248	0.02260
$\lambda_{11,2}$	-0.2	0.00013	0.02256	-0.00204	0.02572	-0.00096	0.02267	-0.00259	0.02628	-0.00122	0.02354
ρ_{11}	-0.2	-0.00583	0.06664	0.00098	0.06941	-0.00040	0.07170	0.01529	0.08047	0.01857	0.08446
ρ_{12}	-0.1	-0.00048	0.07936	0.00398	0.08989	0.00403	0.09009	0.01167	0.09889	0.01737	0.09786
c_{11}	1	0.00081	0.02336	-0.00170	0.02603	0.00043	0.02369	-0.00181	0.02624	0.00060	0.02415
c_{12}	1	-0.00103	0.02473	-0.00270	0.02627	-0.00073	0.02483	-0.00270	0.02603	-0.00027	0.02502
c_{13}	1	-0.00064	0.02382	-0.00328	0.02435	-0.00088	0.02387	-0.00385	0.02484	-0.00063	0.02430
Eq. 2	Set II										
b_{12}	0.3	-0.00051	0.01284	0.00368	0.01333	0.00063	0.01272	0.00460	0.01373	0.00183	0.01314
$\lambda_{22,1}$	-0.3	-0.00001	0.02002	-0.00048	0.02121	-0.00035	0.02044	-0.00236	0.02247	-0.00223	0.02137
$\lambda_{22,2}$	-0.15	-0.00011	0.02255	-0.00181	0.02561	-0.00109	0.02320	-0.00095	0.02537	-0.00107	0.02332
$\rho_{2,1}$	-0.1	-0.01172	0.06573	-0.00280	0.07240	-0.00337	0.07011	0.01500	0.07719	0.02104	0.07884
$\rho_{2,2}$	0	-0.00309	0.07527	0.00096	0.08191	0.00067	0.08050	0.01465	0.09191	0.01799	0.08863
c_{24}	1	-0.00137	0.02099	-0.00218	0.02283	-0.00098	0.02083	-0.00211	0.02312	0.00028	0.02141
c_{25}	1	0.00188	0.02360	-0.00094	0.02559	0.00193	0.02378	-0.00080	0.02598	0.00219	0.02341
c_{26}	1	0.00154	0.02132	0.00044	0.02186	0.00186	0.02115	-0.00006	0.02244	0.00263	0.02185
		$n = 1000$									
Eq. 1	Set II										
b_{21}	0.15	-0.00005	0.00963	0.00225	0.01002	0.00095	0.00952	0.00276	0.01010	0.00164	0.00944
$\lambda_{11,1}$	-0.3	-0.00124	0.01310	-0.00185	0.01475	-0.00152	0.01313	-0.00248	0.01496	-0.00200	0.01315
$\lambda_{11,2}$	-0.2	0.00160	0.01311	0.00039	0.01462	0.00053	0.01359	0.00064	0.01485	0.00041	0.01369
ρ_{11}	-0.2	-0.00098	0.04289	0.00198	0.04644	0.00138	0.04717	0.00584	0.05098	0.00817	0.04907
ρ_{12}	-0.1	-0.00279	0.05461	-0.00069	0.06176	-0.00095	0.05952	0.00522	0.06380	0.00545	0.06143
c_{11}	1	-0.00005	0.01546	-0.00106	0.01656	-0.00043	0.01574	-0.00120	0.01689	-0.00002	0.01566
c_{12}	1	-0.00084	0.01666	-0.00189	0.01768	-0.00079	0.01661	-0.00186	0.01789	-0.00063	0.01655
c_{13}	1	-0.00031	0.01654	-0.00131	0.01828	-0.00033	0.01661	-0.00148	0.01816	0.00014	0.01655
Eq. 2	Set II										
b_{12}	0.3	-0.00024	0.00878	0.00135	0.00868	0.00021	0.00876	0.00209	0.00885	0.00079	0.00874
$\lambda_{22,1}$	-0.3	0.00054	0.01225	0.00063	0.01402	0.00008	0.01245	-0.00016	0.01392	-0.00099	0.01262
$\lambda_{22,2}$	-0.15	0.00016	0.01450	-0.00128	0.01557	-0.00078	0.01466	-0.00122	0.01616	-0.00086	0.01457
ρ_{21}	-0.1	-0.00490	0.04463	-0.00314	0.04672	-0.00352	0.04549	0.00840	0.04993	0.00925	0.04872
ρ_{22}	0	-0.00564	0.05227	-0.00058	0.05685	-0.00180	0.05542	0.00351	0.06174	0.00394	0.06008
c_{24}	1	0.00065	0.01635	-0.00079	0.01796	0.00075	0.01611	-0.00106	0.01797	0.00128	0.01631
c_{25}	1	-0.00005	0.01689	-0.00066	0.01835	0.00042	0.01672	-0.00066	0.01816	0.00035	0.01691
c_{26}	1	-0.00020	0.01705	-0.00054	0.01800	0.00040	0.01720	-0.00040	0.01794	0.00008	0.01758

Table 6: Bias and RMSE of Various Estimators for Parameters of Equation 1 and 2, Parameter Set III with Social Interaction Weights Matrices

		Soc. Interact. Weights Matrices; $n = 500$									
Parameter	True	MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	Set III	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	0.00051	0.01351	0.00411	0.01367	0.00153	0.01336	0.00532	0.01429	0.00296	0.01359
$\lambda_{11,1}$	0	0.00064	0.02031	0.00157	0.02164	0.00258	0.01980	-0.00137	0.02182	-0.00065	0.02038
$\lambda_{11,2}$	0	-0.00035	0.02179	-0.00331	0.02369	-0.00211	0.02164	-0.00256	0.02472	-0.00116	0.02210
ρ_{11}	0	-0.00538	0.06453	-0.00145	0.06674	-0.00351	0.06660	0.02438	0.07803	0.03037	0.08000
ρ_{12}	0	-0.00066	0.07220	0.00343	0.08182	0.00414	0.08272	0.01411	0.09080	0.01733	0.08996
c_{11}	1	-0.00008	0.02398	-0.00109	0.02643	-0.00098	0.02389	-0.00140	0.02655	0.00016	0.02400
c_{12}	1	-0.00064	0.02561	-0.00349	0.02604	-0.00152	0.02548	-0.00307	0.02642	-0.00157	0.02532
c_{13}	1	0.00004	0.02425	-0.00380	0.02496	-0.00081	0.02396	-0.00412	0.02459	0.00003	0.02428
Eq. 2	Set III										
b_{12}	0.3	-0.00058	0.01339	0.00296	0.01351	-0.00019	0.01331	0.00444	0.01403	0.00156	0.01321
$\lambda_{22,1}$	0	0.00022	0.01861	0.00168	0.02032	0.00204	0.01930	-0.00131	0.02129	-0.00100	0.01936
$\lambda_{22,2}$	0	-0.00003	0.02049	-0.00322	0.02294	-0.00204	0.02120	-0.00174	0.02329	-0.00087	0.02135
ρ_{21}	0	-0.01218	0.06197	-0.00768	0.06843	-0.00745	0.06707	0.01536	0.07472	0.02470	0.07542
ρ_{22}	0	-0.00391	0.07103	0.00116	0.07894	0.00047	0.07619	0.01238	0.08824	0.01798	0.08705
c_{24}	1	-0.00168	0.02193	-0.00191	0.02265	-0.00146	0.02148	-0.00153	0.02323	-0.00036	0.02207
c_{25}	1	0.00202	0.02401	-0.00112	0.02574	0.00178	0.02385	-0.00108	0.02594	0.00186	0.02404
c_{26}	1	0.00165	0.02223	0.00001	0.02288	0.00163	0.02225	0.00002	0.02297	0.00251	0.02240
		$n = 1000$									
Eq. 1	Set III										
b_{21}	0.15	-0.00019	0.00980	0.00216	0.01014	0.00073	0.00982	0.00270	0.01020	0.00148	0.00983
$\lambda_{11,1}$	0	-0.00137	0.01282	-0.00125	0.01434	-0.00053	0.01274	-0.00243	0.01458	-0.00166	0.01262
$\lambda_{11,2}$	0	0.00117	0.01311	0.00020	0.01429	0.00034	0.01296	0.00049	0.01457	0.00066	0.01320
ρ_{11}	0	-0.00084	0.04160	0.00055	0.04401	-0.00010	0.04533	0.01129	0.04862	0.01391	0.04871
ρ_{12}	0	-0.00404	0.05041	-0.00108	0.05518	-0.00073	0.05425	0.00526	0.05768	0.00658	0.05572
c_{11}	1	-0.00017	0.01602	-0.00130	0.01694	-0.00051	0.01604	-0.00150	0.01737	-0.00018	0.01606
c_{12}	1	-0.00070	0.01680	-0.00204	0.01831	-0.00084	0.01676	-0.00202	0.01841	-0.00055	0.01661
c_{13}	1	-0.00007	0.01704	-0.00142	0.01850	-0.00001	0.01701	-0.00155	0.01874	-0.00004	0.01669
Eq. 2	Set III										
b_{12}	0.3	-0.00006	0.00933	0.00141	0.00931	0.00051	0.00958	0.00216	0.00943	0.00102	0.00936
$\lambda_{22,1}$	0	0.00036	0.01268	0.00117	0.01364	0.00112	0.01266	0.00000	0.01386	-0.00014	0.01250
$\lambda_{22,2}$	0	-0.00027	0.01352	-0.00190	0.01480	-0.00093	0.01373	-0.00146	0.01491	-0.00062	0.01395
ρ_{21}	0	-0.00587	0.04305	-0.00522	0.04489	-0.00467	0.04487	0.00870	0.04900	0.01195	0.04819
ρ_{22}	0	-0.00412	0.04906	-0.00045	0.05481	-0.00189	0.05234	0.00245	0.05758	0.00334	0.05741
c_{24}	1	0.00060	0.01659	-0.00079	0.01826	0.00057	0.01656	-0.00093	0.01864	0.00076	0.01653
c_{25}	1	-0.00011	0.01692	-0.00058	0.01836	0.00026	0.01674	-0.00060	0.01811	0.00036	0.01664
c_{26}	1	-0.00019	0.01755	-0.00100	0.01838	-0.00011	0.01741	-0.00057	0.01812	0.00001	0.01753

Table 7: Bias and RMSE of Various Estimators for Parameters of Equation 1 and 2, Parameter Set I with Spatial Rook-type Weights Matrices

		Rook-type Weights Matrices; $n = 484$									
Parameter	True	MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	Set I	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	0.00030	0.01395	0.00435	0.01465	0.00117	0.01405	0.00594	0.01540	0.00283	0.01471
$\lambda_{11,1}$	0.3	0.00011	0.01822	-0.00208	0.02199	-0.00112	0.01904	-0.00171	0.02134	-0.00096	0.01863
$\lambda_{11,2}$	0.2	-0.00027	0.01813	-0.00202	0.02055	0.00016	0.01813	-0.00390	0.02091	-0.00242	0.01926
ρ_{11}	0.2	-0.00284	0.05874	-0.00119	0.06197	-0.00082	0.06418	0.00668	0.06494	0.00552	0.06381
ρ_{12}	0.1	-0.00731	0.05813	-0.00547	0.06625	-0.00582	0.06349	0.01667	0.07192	0.02253	0.06910
c_{11}	1	0.00017	0.02225	-0.00171	0.02517	-0.00029	0.02214	-0.00215	0.02463	-0.00042	0.02203
c_{12}	1	-0.00092	0.02288	-0.00353	0.02509	-0.00174	0.02278	-0.00361	0.02606	-0.00216	0.02249
c_{13}	1	0.00039	0.02429	-0.00124	0.02531	-0.00008	0.02425	-0.00181	0.02564	-0.00018	0.02391
Eq. 2	Set I										
b_{12}	0.3	0.00042	0.01300	0.00345	0.01401	0.00044	0.01340	0.00460	0.01420	0.00151	0.01346
$\lambda_{22,1}$	0.3	-0.00097	0.01852	-0.00265	0.01951	-0.00244	0.01884	-0.00188	0.01960	-0.00111	0.01854
$\lambda_{22,2}$	0.15	0.00026	0.01878	0.00029	0.02081	0.00137	0.01903	-0.00228	0.02195	-0.00142	0.01938
ρ_{21}	0.1	0.00073	0.06013	0.00307	0.06813	0.00118	0.06898	0.00316	0.07124	0.00520	0.07404
ρ_{22}	0	-0.00393	0.05632	-0.00543	0.06503	-0.00569	0.06153	0.01548	0.07350	0.01855	0.07269
c_{24}	1	-0.00136	0.02601	-0.00336	0.02745	-0.00211	0.02571	-0.00275	0.02854	-0.00005	0.02615
c_{25}	1	0.00135	0.02378	0.00018	0.02529	0.00165	0.02374	0.00107	0.02505	0.00284	0.02382
c_{26}	1	0.00064	0.02272	-0.00100	0.02444	0.00089	0.02278	-0.00088	0.02485	0.00135	0.02316
		$n = 961$									
Eq. 1	Set I										
b_{21}	0.15	-0.00020	0.00994	0.00213	0.00997	0.00057	0.00985	0.00291	0.01039	0.00137	0.01003
$\lambda_{11,1}$	0.3	0.00011	0.01352	-0.00205	0.01516	-0.00111	0.01333	-0.00200	0.01520	-0.00122	0.01369
$\lambda_{11,2}$	0.2	-0.00049	0.01325	0.00018	0.01531	0.00015	0.01369	-0.00088	0.01560	-0.00093	0.01364
ρ_{11}	0.2	-0.00232	0.04129	0.00099	0.04607	0.00009	0.04424	0.00234	0.04798	0.00135	0.04576
ρ_{12}	0.1	-0.00231	0.04338	-0.00043	0.04829	-0.00213	0.04669	0.01064	0.05075	0.01222	0.04986
c_{11}	1	0.00047	0.01766	-0.00118	0.02092	-0.00007	0.01789	-0.00133	0.02058	0.00003	0.01759
c_{12}	1	-0.00015	0.01756	-0.00107	0.01857	-0.00093	0.01723	-0.00121	0.01807	-0.00072	0.01741
c_{13}	1	0.00018	0.01677	-0.00100	0.01879	-0.00053	0.01696	-0.00112	0.01907	-0.00025	0.01729
Eq. 2	Set I										
b_{12}	0.3	-0.00005	0.00988	0.00176	0.00988	0.00010	0.00984	0.00223	0.00991	0.00048	0.00970
$\lambda_{22,1}$	0.3	0.00083	0.01260	-0.00072	0.01500	0.00027	0.01288	-0.00035	0.01545	0.00047	0.01291
$\lambda_{22,2}$	0.15	-0.00085	0.01385	-0.00089	0.01509	-0.00003	0.01399	-0.00129	0.01495	-0.00133	0.01355
ρ_{21}	0.1	-0.00295	0.04679	-0.00214	0.05208	-0.00179	0.05039	0.00034	0.05322	0.00105	0.05215
ρ_{22}	0	-0.00502	0.04425	-0.00498	0.04669	-0.00673	0.04552	0.00361	0.04921	0.00553	0.04831
c_{24}	1	-0.00025	0.01627	-0.00113	0.01750	-0.00014	0.01633	-0.00111	0.01757	0.00046	0.01627
c_{25}	1	0.00193	0.01854	0.00104	0.02031	0.00199	0.01834	0.00108	0.02028	0.00236	0.01872
c_{26}	1	-0.00002	0.01765	-0.00024	0.01862	0.00012	0.01764	-0.00041	0.01872	0.00056	0.01791

Table 8: Bias and RMSE of Various Estimators for Parameters of Equation 1 and 2, Parameter Set II with Spatial Rook-type Weights Matrices

		Rook-type Weights Matrices; $n = 484$									
Parameter	True	MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	Set II	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	-0.00006	0.01373	0.00475	0.01461	0.00228	0.01399	0.00584	0.01461	0.00308	0.01421
$\lambda_{11,1}$	-0.3	0.00045	0.01847	-0.00196	0.02171	-0.00230	0.01901	0.00049	0.02111	-0.00015	0.01959
$\lambda_{11,2}$	-0.2	-0.00055	0.01779	-0.00040	0.02000	-0.00030	0.01798	-0.00360	0.02016	-0.00253	0.01860
ρ_{11}	-0.2	-0.00300	0.06501	0.00157	0.06740	0.00245	0.06735	-0.00580	0.07495	-0.00761	0.07299
ρ_{12}	-0.1	-0.00452	0.05882	-0.00295	0.06462	-0.00408	0.06426	0.01089	0.07022	0.01641	0.06848
c_{11}	1	0.00038	0.02213	-0.00269	0.02446	-0.00012	0.02218	-0.00282	0.02527	0.00036	0.02140
c_{12}	1	-0.00033	0.02182	-0.00202	0.02491	-0.00088	0.02176	-0.00321	0.02532	-0.00127	0.02231
c_{13}	1	-0.00001	0.02382	-0.00076	0.02489	-0.00008	0.02342	-0.00137	0.02492	0.00038	0.02388
Eq. 2	Set II										
b_{12}	0.3	0.00038	0.01337	0.00382	0.01357	0.00132	0.01304	0.00486	0.01361	0.00237	0.01324
$\lambda_{22,1}$	-0.3	-0.00028	0.01615	-0.00384	0.01746	-0.00352	0.01687	-0.00144	0.01812	-0.00132	0.01735
$\lambda_{22,2}$	-0.15	0.00040	0.01871	0.00151	0.02009	0.00093	0.01886	-0.00270	0.02092	-0.00204	0.01890
ρ_{21}	-0.1	0.00231	0.05824	0.00559	0.06581	0.00496	0.06723	-0.00082	0.07060	-0.00308	0.06936
ρ_{22}	0	-0.00225	0.05764	-0.00465	0.06666	-0.00518	0.06369	0.01972	0.07576	0.02030	0.07460
c_{24}	1	-0.00167	0.02478	-0.00270	0.02770	-0.00118	0.02492	-0.00279	0.02755	-0.00089	0.02507
c_{25}	1	0.00167	0.02311	0.00043	0.02466	0.00247	0.02307	0.00132	0.02549	0.00372	0.02373
c_{26}	1	0.00043	0.02266	-0.00114	0.02475	0.00133	0.02278	-0.00050	0.02506	0.00165	0.02258
		$n = 961$									
Eq. 1	Set II										
b_{21}	0.15	-0.00011	0.00966	0.00263	0.01011	0.00133	0.00982	0.00290	0.01024	0.00169	0.01006
$\lambda_{11,1}$	-0.3	-0.00040	0.01361	-0.00183	0.01531	-0.00231	0.01383	-0.00110	0.01557	-0.00106	0.01378
$\lambda_{11,2}$	-0.2	0.00017	0.01322	0.00040	0.01476	0.00078	0.01342	-0.00052	0.01499	-0.00063	0.01366
ρ_{11}	-0.2	-0.00160	0.04336	0.00207	0.04970	0.00114	0.04733	-0.00482	0.05256	-0.00571	0.05156
ρ_{12}	-0.1	-0.00155	0.04317	-0.00096	0.04821	-0.00084	0.04759	0.00781	0.05042	0.00993	0.05103
c_{11}	1	0.00062	0.01648	-0.00116	0.01963	0.00056	0.01658	-0.00114	0.01942	0.00073	0.01659
c_{12}	1	-0.00061	0.01827	-0.00117	0.01840	-0.00084	0.01825	-0.00165	0.01857	-0.00110	0.01856
c_{13}	1	-0.00018	0.01717	-0.00169	0.01856	-0.00018	0.01727	-0.00194	0.01857	-0.00053	0.01708
Eq. 2	Set II										
b_{12}	0.3	-0.00007	0.01019	0.00205	0.01025	0.00066	0.00994	0.00230	0.01002	0.00109	0.00978
$\lambda_{22,1}$	-0.3	0.00016	0.01393	-0.00141	0.01486	-0.00138	0.01383	-0.00070	0.01518	-0.00085	0.01406
$\lambda_{22,2}$	-0.15	-0.00072	0.01295	0.00010	0.01378	0.00011	0.01288	-0.00118	0.01409	-0.00128	0.01283
ρ_{21}	-0.1	-0.00120	0.04742	0.00031	0.05182	-0.00029	0.04901	-0.00144	0.05227	-0.00320	0.05098
ρ_{22}	0	-0.00526	0.04545	-0.00502	0.04915	-0.00615	0.04796	0.00688	0.05104	0.00824	0.05135
c_{24}	1	-0.00038	0.01584	-0.00090	0.01691	0.00004	0.01600	-0.00119	0.01723	0.00035	0.01593
c_{25}	1	0.00232	0.01822	0.00116	0.01978	0.00255	0.01816	0.00112	0.02001	0.00274	0.01842
c_{26}	1	0.00011	0.01678	-0.00064	0.01853	0.00034	0.01653	-0.00092	0.01872	0.00069	0.01707

Table 9: Bias and RMSE of Various Estimators for Parameters of Equation 1 and 2, Parameter Set III with Spatial Rook-type Weights Matrices

		Rook-type Weights Matrices; $n = 484$									
Parameter	True	MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	Set III	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	0.00006	0.01409	0.00481	0.01473	0.00164	0.01413	0.00604	0.01520	0.00287	0.01447
$\lambda_{11,1}$	0	0.00049	0.01856	-0.00249	0.02100	-0.00185	0.01905	-0.00088	0.02125	-0.00081	0.01923
$\lambda_{11,2}$	0	-0.00052	0.01833	-0.00103	0.02026	-0.00008	0.01808	-0.00350	0.02026	-0.00270	0.01909
ρ_{11}	0	-0.00328	0.06282	0.00100	0.06496	-0.00031	0.06585	0.00211	0.06942	0.00081	0.06853
ρ_{12}	0	-0.00586	0.05788	-0.00460	0.06559	-0.00365	0.06381	0.01654	0.07267	0.02004	0.07024
c_{11}	1	0.00052	0.02259	-0.00258	0.02475	-0.00010	0.02211	-0.00284	0.02491	-0.00009	0.02209
c_{12}	1	-0.00062	0.02244	-0.00276	0.02536	-0.00166	0.02224	-0.00340	0.02553	-0.00159	0.02303
c_{13}	1	0.00020	0.02385	-0.00107	0.02594	-0.00019	0.02377	-0.00170	0.02576	0.00023	0.02408
Eq. 2	Set III										
b_{12}	0.3	0.00042	0.01381	0.00390	0.01438	0.00114	0.01395	0.00500	0.01457	0.00209	0.01349
$\lambda_{22,1}$	0	-0.00073	0.01763	-0.00285	0.01822	-0.00266	0.01756	-0.00146	0.01870	-0.00087	0.01763
$\lambda_{22,2}$	0	0.00010	0.01939	0.00088	0.02074	0.00149	0.01963	-0.00256	0.02204	-0.00164	0.01976
ρ_{21}	0	0.00189	0.05972	0.00508	0.06761	0.00385	0.06899	0.00152	0.07181	0.00129	0.07172
ρ_{22}	0	-0.00346	0.05684	-0.00583	0.06572	-0.00563	0.06269	0.01748	0.07473	0.01955	0.07454
c_{24}	1	-0.00182	0.02561	-0.00292	0.02810	-0.00190	0.02554	-0.00312	0.02827	-0.00044	0.02600
c_{25}	1	0.00155	0.02290	0.00019	0.02515	0.00217	0.02307	0.00117	0.02570	0.00333	0.02380
c_{26}	1	0.00077	0.02280	-0.00111	0.02489	0.00104	0.02282	-0.00070	0.02444	0.00169	0.02290
		$n = 961$									
Eq. 1	Set III										
b_{21}	0.15	-0.00014	0.00983	0.00254	0.01020	0.00094	0.00975	0.00298	0.01036	0.00156	0.01015
$\lambda_{11,1}$	0	0.00002	0.01279	-0.00200	0.01454	-0.00135	0.01310	-0.00156	0.01482	-0.00097	0.01308
$\lambda_{11,2}$	0	-0.00005	0.01349	0.00046	0.01501	0.00091	0.01360	-0.00090	0.01537	-0.00037	0.01352
ρ_{11}	0	-0.00247	0.04337	0.00111	0.04784	0.00067	0.04729	-0.00025	0.05003	-0.00013	0.04842
ρ_{12}	0	-0.00140	0.04384	-0.00116	0.04852	-0.00188	0.04792	0.00981	0.05217	0.01188	0.05151
c_{11}	1	0.00044	0.01753	-0.00116	0.02016	0.00018	0.01746	-0.00118	0.02024	0.00041	0.01719
c_{12}	1	-0.00037	0.01807	-0.00074	0.01888	-0.00071	0.01797	-0.00151	0.01862	-0.00108	0.01811
c_{13}	1	0.00015	0.01709	-0.00125	0.01883	-0.00018	0.01698	-0.00128	0.01906	-0.00040	0.01712
Eq. 2	Set III										
b_{12}	0.3	0.00017	0.01017	0.00186	0.01019	0.00059	0.00994	0.00238	0.01023	0.00086	0.00994
$\lambda_{22,1}$	0	0.00010	0.01292	-0.00102	0.01478	-0.00069	0.01295	-0.00045	0.01513	-0.00065	0.01288
$\lambda_{22,2}$	0	-0.00076	0.01333	-0.00014	0.01479	0.00005	0.01361	-0.00137	0.01452	-0.00120	0.01347
ρ_{21}	0	-0.00222	0.04786	-0.00085	0.05181	-0.00219	0.04922	-0.00087	0.05351	-0.00023	0.05171
ρ_{22}	0	-0.00493	0.04463	-0.00465	0.04706	-0.00625	0.04630	0.00571	0.04953	0.00730	0.05053
c_{24}	1	-0.00011	0.01627	-0.00134	0.01729	0.00019	0.01625	-0.00125	0.01746	0.00067	0.01606
c_{25}	1	0.00231	0.01865	0.00101	0.02021	0.00271	0.01862	0.00107	0.02007	0.00250	0.01910
c_{26}	1	0.00030	0.01714	-0.00041	0.01869	0.00029	0.01736	-0.00069	0.01880	0.00062	0.01717

G.3 Monte Carlo Simulations for the Weakly Identified Case

In Appendix F included in this Online Supplementary Appendix we discussed three scenarios where identification from the linear moment conditions fails. In this case the two step GS2SLS and GS3SLS estimators, which rely on the linear moment conditions, will be inconsistent. In the following we illustratively report on the performance of ML, and one-step and two-step estimators for scenarios that are "close" to the non-identified scenarios. We refer to those experiments as Weak Identification Scenario 1, 2 and 3.

■ Weak Identification Scenario 1

To explore this scenario we consider a simplified variant of the first equation of the model discussed in Appendix G.1, based on parameter Set I, but with $b_{21} = \lambda_{11,2} = \rho_{11} = \rho_{12} = 0$. To explore the effect of the X instruments being weak we consider a case where the parameters on the exogenous variables in equation 1 are $c_{1k} = 0.0001$, $k = 1, 2, 3$. The simulations employ the social interaction weights matrix $\mathbf{M}_1$ with sample size $n = 500$. The results reported in Table 10 at the end of this subsection show that in this case the GS2SLS estimator for autoregressive parameter $\lambda_{11,1}$ is substantially biased, in contrast to ML and LQ-GS2SLS.

■ Weak Identification for Scenario 2

To explore the effect when the weights matrix is close to the case where all weights are equal (except for the zero diagonal elements) we consider a simplified variant of the first equation of the model discussed in Appendix G.1, based on parameter Set I, but with $b_{21} = \lambda_{11,2} = \rho_{11} = \rho_{12} = 0$. For this simulation the social interaction weights matrix $\mathbf{M}_1$ with sample size $n = 500$ is generated as in Appendix G.1, however with $d_* = 1.5$ (instead of $d_* = .3$). This leaves most of the 500 considered individuals with more than 400 friends in the simulations presented in Table 11 at the end of this subsection. As expected, all estimators show bias for the autoregressive parameter $\lambda_{11,1}$. The absolute bias for the ML is close to 0.1. All other estimators exhibit even larger biases.

■ Weak Identification for Scenario 3

To explore this scenario we consider the single equation case with the data generated from the following model:

$$\mathbf{y}_1 = \lambda_{11,1}\mathbf{M}_1\mathbf{y}_1 + \textstyle\sum_{k=1}^{3} c_{k1}\mathbf{x}_k + \sum_{k=1}^{3} c_{k+3,1}\mathbf{M}_1\mathbf{x}_k + \mathbf{u}_1,$$
$$\mathbf{u}_1 = \rho_{11}\mathbf{M}_1\mathbf{u}_1 + \boldsymbol{\epsilon}_1.$$

The matrix $\mathbf{M}_1$, the exogenous covariates $\mathbf{x}_k$ and corresponding parameters c_{k1}, $k = 1, 2, 3$, and the innovations $\boldsymbol{\epsilon}_1$ are as described in Appendix G.1. For the case where $c_{k+3,1} = -\lambda_{11,1}c_{k1}$ and $\rho_{11} = -\lambda_{11,1}$ the above model is approximately equal to the following model without network interdependencies:

$$\mathbf{y}_1 = \sum_{k=1}^{3} c_{k1}\mathbf{x}_k + \boldsymbol{\epsilon}_1.$$

This is readily seen upon pre-multiplying the last equation with $\mathbf{I}-\lambda_{11,1}\mathbf{M}_1$ and exploiting the approximation $\mathbf{u}_1 = (\mathbf{I} - \rho_{11}\mathbf{M}_1)^{-1}\boldsymbol{\epsilon}_1 \doteq \boldsymbol{\epsilon}_1 + \rho_{11}\mathbf{M}_1\boldsymbol{\epsilon}_1$ for the disturbance process. Now let $c_{k+3,1} = -(\lambda_{11,1} + \nu)c_{k1}$ and $\rho_1 = -(\lambda_{11,1} + \nu)$, then we expect the parameters to only be weakly identified for ν small. (Of course, proximity to the non-identified set of parameters can be modeled in various ways. We adopted this approach for its simplicity.) In Table 12 at the end of this subsection we report on the small sample behavior of various estimators for $\lambda_{11,1} = .5$ and $\nu = .3$. The results in Table 12 indicate that for $v = .3$ the ML estimator, and to a lesser degree the LQ-GS2SLS estimator, are still able to estimate the parameters $\lambda_{11,1}$, ρ_{11}, c_{41}, c_{51}, c_{61} reasonably well, while GS2SLS is already severely biased.

Table 10: Bias and RMSE under Scenario 1 of Various Estimators for Parameters of a Simplified Version of Equation 1. Parameter Set I, except that $b_{21} = \lambda_{11,2} = \rho_{11} = \rho_{12} = 0$ and $c_{1k} = 0.0001$ for all $k = \{1, 2, 3\}$.

Parameter		Soc. Interact. Matrix; $n = 500$					
		MLE		GS2SLS		LQ-GS2SLS	
One equ.	True	Bias	RMSE	Bias	RMSE	Bias	RMSE
$\lambda_{11,1}$	0.30	-0.00134	0.04676	0.17807	0.50151	0.00179	0.04714
c_{11}	0.0001	0.00013	0.02149	0.00182	0.02158	0.00034	0.02149
c_{12}	0.0001	-0.00050	0.02439	-0.00090	0.02529	-0.00064	0.02449
c_{13}	0.0001	0.00015	0.02417	0.00058	0.02314	0.00049	0.02410

Table 11: Bias and RMSE under Scenario 2 of Various Estimators for Parameters of a Simplified Version of Equation 1. Parameter Set I, except that $b_{21} = \lambda_{11,2} = \rho_{11} = \rho_{12} = 0$. Social Interaction Weight Matrix with High Number of Friends

Parameter		Soc. Interact. Weights Matrices; $n = 500$									
		MLE		GS2SLS		GS3SLS		LQ-GS2SLS		LQ-GS3SLS	
Eq. 1	True	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
b_{21}	0.15	-0.00009	0.01357	0.00273	0.01387	0.00143	0.01383	0.00316	0.01389	0.00198	0.01378
$\lambda_{11,1}$	0.30	-0.09312	0.44336	0.20473	0.36126	0.20307	0.36233	0.23556	0.43856	0.22588	0.47735
c_{11}	1	0.00037	0.02424	-0.00013	0.02642	0.00076	0.02461	-0.00033	0.02644	0.00084	0.02434
c_{12}	1	-0.00113	0.02546	-0.00142	0.02556	-0.00064	0.02534	-0.00125	0.02591	-0.00004	0.02560
c_{13}	1	-0.00071	0.02450	-0.00155	0.02584	0.00001	0.02476	-0.00189	0.02559	-0.00003	0.02482

Table 12: Bias and RMSE under Scenario 3 of Various Estimators in Neighborhood of Parameter Singularity (where Order of Spillovers is Overspecified)

Parameter		Soc. Interact. Matrix; $n = 500$					
		MLE		GS2SLS		LQ-GS2SLS	
One equ.	True	Bias	RMSE	Bias	RMSE	Bias	RMSE
$\lambda_{11,1}$	0.3	-0.03320	0.17906	-0.34951	0.69577	0.06243	0.20602
ρ_{11}	-0.3	0.03192	0.21385	0.36910	0.55993	-0.08737	0.27132
c_{11}	1	-0.00103	0.03952	-0.00992	0.04777	0.00161	0.03982
c_{12}	1	-0.00511	0.04634	-0.01437	0.05319	-0.00274	0.04586
c_{13}	1	-0.00280	0.04570	-0.01507	0.05369	0.00051	0.04810
c_{14}	-0.4	0.03411	0.18339	0.30034	0.62913	-0.04263	0.18271
c_{15}	-0.4	0.04620	0.18502	0.29730	0.64222	-0.04365	0.18606
c_{16}	-0.4	0.02592	0.16655	0.28633	0.63019	-0.05244	0.19939

References

[1] Bramoullé, Y., H. Djebbari, and B. Fortin (2009) Identification of peer effects through social networks. *Journal of Econometrics* 150, 41-55.

[2] Kelejian, H.H., and I.R. Prucha (2002) 2SLS and OLS in a spatial autoregressive model with equal spatial weights. *Regional Science and Urban Economics* 32, 691-707.

[3] Kelejian, H.H., I.R. Prucha, and Y. Yuzefovich (2006) Estimation problems in models with spatial weighting matrices which have blocks of equal elements. *Journal of Regional Science* 46, 507-517.

[4] Paula, Á. (2017) Econometrics of network models. In B. Honoré, A. Pakes, M. Piazzesi, & L. Samuelson (eds.), *Advances in Economics and Econometrics: Eleventh World Congress (Econometric Society Monographs*, pp. 268-323). Cambridge University Press.