Refined GMM Estimators for Simultaneous Equations Models with Network Interactions

Peter H. Egger[*] and Ingmar R. Prucha[†]

March 7, 2023

Abstract

The paper proposes a refinement of the generalized spatial two-stage and three-stage least squares estimators for simultaneous systems of equations with network interdependence, recently introduced in Drukker, Egger and Prucha (2022). Specifically, we propose a refined weighting of the moment conditions underlying those estimators. Monte Carlo simulations document that the refined weighting potentially achieves non-trivial reductions in the root mean-squared errors for the network parameters of interest.

Key Words: Cliff-Ord spatial model; Two-stage least squares estimation; Three-stage least squares estimation; Generalized method of moments estimation

JEL Codes: C21, C31, C36

Acknowledgement: The authors gratefully acknowledge the helpful comments by an editor and an anonymous reviewer.

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

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

1 Introduction

In this paper we explore modifications to the class of GMM estimators for simultaneous equation systems for cross sectional data with network structures recently introduced in Drukker, Egger and Prucha (2022), henceforth DEP. The aim of those modifications is to improve on the small sample properties of DEP's one-step GMM estimators, which are based on both linear and quadratic moment conditions and referred to as one-step LQ-GS2SLS and LQ-GS3SLS estimators.

The econometric analysis of spatial networks has a long history in geography, regional science and urban economics; see, e.g., Anselin (1988) and Cliff and Ord (1973). Since the mid-1990s the development of econometric methods of inference for Anselin-Cliff-Ord-type models has also been an active area of research in economics, recognizing that possible applications include both spatial and social network structures. Originally the focus of this research was on single equation models. In economics it is common for outcomes to be generated by a system of equations. Kelejian and Prucha (2004) provide an early development of generalized method of moments (GMM) estimators for such models. Recent contributions include Baltagi and Deng (2015), Yang and Lee (2019), and Liu (2020).[1]

DEP develop GMM estimation methodologies for simultaneous equation models that allow for spatial interactions in terms of higher order spatial lags of the endogenous variables, the exogenous variables as well as the disturbances. The aim of this paper is to explore a refinement of the one-step LQ-GS2SLS and LQ-GS3SLS estimators introduced in DEP, with the aim of achieving improvements of their small sample performance.

In more detail, the first order condition of the one-step LQ-GS2SLS and LQ-GS3SLS estimators weighs, in line with optimally weighted GMM estimators, the sample moment vector with the matrix of first-order derivatives of the moment vector normalized by the inverse variance covariance matrix of the sample moment vector. The asymptotic properties of the estimator depend on the probability limit of the matrix of first-order derivatives of the moment vector. An inspection of the first-order derivatives of the moment vector reveals that several quantities converge in probability to zero. We modify the first order conditions by setting those quantities to their probability limit, i.e., to zero. Intuition suggests that by explicitly incorporating available information of the probability limit of those quantities, there may be an improvement in the small sample behavior of the correspondingly defined GMM estimators. The results of our Monte Carlo (MC) simulations are in support of this conjecture. Improvements become especially pronounced when identification is weakened.

[1] See, e.g., DEP for a more extensive review of the literature.

2 Model

The model is assumed to be identical to the simultaneous system of G equations for G endogenous variables observed for n cross sectional units as specified in Section 2 of DEP. The model allows for higher order network dependence in the form of spatial lags in the endogenous variables, exogenous variables and disturbances. To conserve space we only provide here a basic description of the model, and refer the reader to DEP for details. In that paper we also discuss the wider applicability of the model to not only spatial networks, but also social networks, while we continue using the terminology common for spatial models.

The g-th structural equation of the system is given by

$$y_g = Z_g \delta_g + u_g, \tag{1}$$

where y_g denotes the $n \times 1$ vector of observations on the g-th endogenous variable, $Z_g = [Y_g, X_g, \overline{Y}_g]$ denotes the matrix of observations on the covariates appearing in the g-th equation, where Y_g is the matrix of r.h.s. endogenous variables, X_g the matrix of exogenous variables (which may include spatial lags of exogenous variables), $\overline{Y}_g$ denotes the matrix of observations on spatial lags of endogenous variables appearing on the r.h.s. of the g-th equation, δ_g denotes the vector of structural parameters corresponding to Z_g, and u_g denotes the vector of disturbances of the g-th equation. The specification of $\overline{Y}_g$ is general and may contain higher order as well as cross equation spatial lags.

The disturbance process is given by

$$u_g = R_g^*(\rho_g) u_g + \varepsilon_g, \quad \text{with} \quad R_{g,n}^*(\rho_g) = \textstyle\sum_{r \in \mathbf{I}_{g,\rho}} \rho_{g,r} M_r, \tag{2}$$

where $M_1, \ldots, M_q$ denotes the set of all spatial weight matrices appearing in the disturbance processes of all G equations, $I_{g,\rho} \subset \{1, \ldots, q\}$ denotes the set of indexes of the weight matrices appearing in the disturbance process of the g-th equation, ρ_g denotes the vector of corresponding autoregressive parameters, and ε_g denotes the vector of basic innovations.

For purposes of estimation it proves helpful to apply a spatial Cochrane-Orcutt transformation to the model. In particular, premultiplying (1) by $I_n - R_g^*(\rho_g)$ yields

$$y_{*g} = Z_{*g} \delta_g + \varepsilon_g, \tag{3}$$

with $y_{*g} = y_{*g}(\rho_g) = \left[I_n - R_g^*(\rho_g)\right] y_g$, $Z_{*g} = Z_{*g}(\rho_g) = \left[I_n - R_g^*(\rho_g)\right] Z_g$. Stacking the transformed equations yields

$$y_* = Z_* \delta + \varepsilon, \tag{4}$$

with $y_* = [y_{*1}', \ldots, y_{*G}']'$, $Z_* = diag_{g=1}^G [Z_{*g}]$, $\delta = [\delta_1', \ldots, \delta_G']'$, and $\varepsilon = [\varepsilon_1', \ldots, \varepsilon_G']'$. We maintain the set of assumptions as given in DEP. Under those assumptions $E\varepsilon = 0$, $E\varepsilon\varepsilon' = \Sigma \otimes I_n$ with $\Sigma = (\sigma_{ij})_{i,j=1,\ldots,G}$ non-singular.

3 GMM Estimators

We next give a brief review of the one-step LQ-GS2SLS and LQ-GS3SLS estimators introduced in DEP. Let $\theta_g = [\rho_g', \delta_g']'$, $\theta = [\theta_1', \ldots, \theta_G']'$ and let $\theta_{0,g} = [\rho_{0,g}', \delta_{0,g}']'$, $\theta_0 = [\theta_{0,1}', \ldots, \theta_{0,G}']'$ denote the true parameters. From (3) we have

$$\varepsilon_g = \varepsilon_g(\theta_{0,g}) = \left[I_n - R_g^*(\rho_{0,g})\right][y_g - Z_g\delta_{0,g}].$$

The one-step LQ-GS2SLS and LQ-GS3SLS are then based on the following vectors of linear and quadratic sample moments ($g = 1, \ldots, G$):

$$m_g(\theta_g) = \begin{bmatrix} m_g^L(\theta_g) \\ m_g^Q(\theta_g) \end{bmatrix},$$

with

$$m_g^L(\theta_g) = n^{-1}H'\varepsilon_g(\theta_g), \qquad m_g^Q(\theta_g) = n^{-1}\begin{bmatrix} \varepsilon_g(\theta_g)'A_1\varepsilon_g(\theta_g) \\ \vdots \\ \varepsilon_g(\theta_g)'A_S\varepsilon_g(\theta_g) \end{bmatrix},$$

where H is a $n \times P$ matrix of instruments and the A_s, $s = 1, \ldots, S$ are $n \times n$ matrices with zero diagonal elements. Specific choices for H and A_s are motivated and described in DEP. The choices for H involve the exogenous variables in the system and spatial lags thereof. The choices for the A_s matrices involve the spatial weight matrices and products of those matrices (with the diagonal elements set to zero).

As shown in DEP we have $Em_g(\theta_g^0) = 0$ and

$$\Phi_{gg} = VC(n^{1/2}m_g(\theta_{0,g})) = \begin{bmatrix} \Psi_{gg}^{LL} & 0 \\ 0 & \Psi_{gg}^{QQ} \end{bmatrix}$$

with $\Psi_{gg}^{LL} = \sigma_{gg}\left[n^{-1}H'H\right]$ and $\Psi_{gg}^{QQ} = \sigma_{gg}^2 K^{QQ}$, where $K^{QQ} = (k_{rs}^{QQ})$ and

$$k_{rs}^{QQ} = (2n)^{-1}tr\left[(A_r + A_r')(A_s + A_s')\right].$$

Let $\widetilde{\Phi}_{gg}$, $\widetilde{\Psi}_{gg}^{LL}$ and $\widetilde{\Psi}_{gg}^{QQ}$ denote the corresponding estimators, where σ_{gg} is replaced by some consistent estimator $\widetilde{\sigma}_{gg}$. Then the one-step LQ-GS2SLS estimator is the limited information GMM estimator defined as

$$\widehat{\theta}_g^o = \arg\min_{\theta_g} \quad m_g(\theta_g)'\widetilde{\Phi}_{gg}^{-1}m(\theta_g). \tag{5}$$

Now consider the stacked sample moment vector

$$m(\theta) = \left[m_1^L(\theta_1)', \ldots, m_G^L(\theta_G)', m_1^Q(\theta_1)', \ldots, m_G^Q(\theta_G)'\right]'.$$

Then, as shown in DEP, we have $Em(\theta^0) = 0$ and

$$\Phi = VC(n^{1/2}m(\theta_0)) = \begin{bmatrix} \Psi^{LL} & 0 \\ 0 & \Psi^{QQ} \end{bmatrix},$$

with

$$\Psi^{LL} = \Sigma \otimes n^{-1} H'H \qquad \text{and} \qquad \Psi^{QQ} = \Sigma_{SQ} \otimes K^{QQ},$$

where $\Sigma_{SQ} = (\sigma^2_{gh})$. Let $\widetilde{\Phi}$, $\widetilde{\Psi}^{LL}$ and $\widetilde{\Psi}^{QQ}$ denote the corresponding estimators. Then the one-step LQ-GS3SLS estimator is the full information GMM estimator defined as

$$\widehat{\widetilde{\theta}}^o = \arg\min_{\theta} \quad m(\theta)'\widetilde{\Phi}^{-1}m(\theta). \tag{6}$$

4 Refined GMM Estimators

Towards defining our refined GMM estimators let

$$G_g^{LL} = \frac{\partial m_g^L(\theta_g)}{\partial \delta_g},\ G_g^{LQ} = \frac{\partial m_g^L(\theta_g)}{\partial \rho_g},\ G_g^{QL} = \frac{\partial m_g^Q(\theta_g)}{\partial \delta_g},\ G_g^{QQ} = \frac{\partial m_g^Q(\theta_g)}{\partial \rho_g}.$$

Then,

$$\begin{aligned} G_g(\theta_g) &= \frac{\partial m_g(\theta_g)}{\partial \theta_g} = \begin{bmatrix} G_g^{LL}(\theta_g) & G_g^{LQ}(\theta_g) \\ G_g^{QL}(\theta_g) & G_g^{QQ}(\theta_g) \end{bmatrix}, \\ G(\theta) &= \frac{\partial m(\theta)}{\partial \theta} = \begin{bmatrix} diag_{g=1}^G[G_g^{LL}(\theta_g)] & diag_{g=1}^G[G_g^{LQ}(\theta_g)] \\ diag_{g=1}^G[G_g^{QL}(\theta_g)] & diag_{g=1}^G[G_g^{QQ}(\theta_g)] \end{bmatrix}, \end{aligned}$$

observing that $\partial m_g/\partial\theta_h = 0$ for $g \neq h$. The one-step LQ-GS2SLS and LQ-GS3SLS solve the following first-order conditions, respectively:

$$G_g(\widehat{\theta}_g^o)'\widetilde{\Phi}_{gg}^{-1}m(\widehat{\theta}_g^o) = 0 \qquad \text{and} \qquad G(\widehat{\widetilde{\theta}}^o)'\widetilde{\Phi}^{-1}m(\widehat{\widetilde{\theta}}^o) = 0. \tag{7}$$

DEP derive the probability limits of the submatrices of $G_g(\theta_g^0)$. In particular they establish that

$$G_g^{LQ}(\theta_{0,g}) = o_p(1) \qquad \text{and} \qquad G_g^{QL}(\theta_{0,g}) = \overline{G}_g^{QL}(\theta_{0,g}) + o_p(1)$$

where

$$G_g^{QL}(\theta_{0,g}) = \begin{bmatrix} -n^{-1}\varepsilon_g'[A_1 + A_1']Z_{*g}(\rho_g) \\ \vdots \\ -n^{-1}\varepsilon_g'[A_S + A_S']Z_{*g}(\rho_g) \end{bmatrix},\ \overline{G}_g^{QL}(\theta_{0,g}) = \begin{bmatrix} -n^{-1}\varepsilon_g'[A_1 + A_1']Z_{*g}^-(\rho_g) \\ \vdots \\ -n^{-1}\varepsilon_g'[A_S + A_S']Z_{*g}^-(\rho_g) \end{bmatrix}$$

with $Z_{*g}^-(\rho_g) = [I_n - R_g^*(\rho_g)Z_g^-$ and $Z_g^- = [Y_g, 0, \overline{Y}_g]$, i.e., where Z_g^- is obtained from $Z_g = [Y_g, X_g, \overline{Y}_g]$ by replacing X_g with a matrix of zeros.

Hence, the derivative matrices $G_g(\theta_g)$ and $G(\theta)$ contain terms that converge to zero. One can conjecture that setting them to their probability limit of zero may be helpful in small samples for estimating θ_g and θ. We next define a refined version of the one-step LQ-GS2SLS and LQ-GS3SLS corresponding to this conjecture. In particular, let

$$\overline{G}_g(\theta_g) = \begin{bmatrix} G_g^{LL}(\theta_g) & 0 \\ \overline{G}_g^{QL}(\theta_g) & G_g^{QQ}(\theta_g) \end{bmatrix} \text{ and } G(\theta) = \begin{bmatrix} diag_{g=1}^G[G_g^{LL}(\theta_g)] & diag_{g=1}^G[0] \\ diag_{g=1}^G[\overline{G}_g^{QL}(\theta_g)] & diag_{g=1}^G[G_g^{QQ}(\theta_g)] \end{bmatrix}.$$

Then, the refined one-step LQ-GS2SLS and LQ-GS3SLS, say $\widehat{\theta}^R_{gg}$ and $\widehat{\widehat{\theta}}^R$, are defined by

$$\overline{G}_g(\widehat{\theta}^R_{gg})'\widetilde{\Phi}^{-1}_{gg}m(\widehat{\theta}^R_{gg}) = 0 \qquad \text{and} \qquad \overline{G}(\widehat{\widehat{\theta}}^R_g)'\widetilde{\Phi}^{-1}m(\widehat{\widehat{\theta}}^R) = 0. \tag{8}$$

5 Monte Carlo Design and Main Results

In what follows we present results from an MC study which assesses the small-sample properties of the LQ-GS2SLS and LQ-GS3SLS of the DEP estimators in comparison to their refined counterparts introduced above.

For the MC study we consider the specialized two-equation system based on (1) and (2) where for $g \in \{1, 2\}$,

$$\begin{aligned} Z_1 &= [y_2, M_1y_1, M_2y_1, x_1, x_2, x_3], \quad \delta_1 = [b_{12}, \lambda_{11}, \lambda_{12}, c_{11}, c_{12}, c_{13}]', \\ Z_2 &= [y_1, M_1y_2, M_2y_2, x_4, x_5, x_6], \quad \delta_2 = [b_{21}, \lambda_{21}, \lambda_{22}, c_{24}, c_{25}, c_{26}]', \\ u_g &= [\rho_{g1}M_1 + \rho M_2]u_g + \varepsilon_g. \end{aligned}$$

Our specification adopts the stylized social-network design of DEP. This design emulates groups of friends in a classroom setting. We refer the reader to DEP for details. Each school has 50 students distributed over three classroom of size 10, 15, and 25, respectively. We generate M_1 and M_2 as social-interactions matrices involving, respectively, closer and less-close friends among class fellows. We conducted MC simulations for different sample sizes and parameter sets. The full set of results for those simulations is made available online. For the results reported below the number of schools is taken to be 2, implying a sample size n of 100. The social interactions matrices M_1 and M_2 are generated outside the MC loop, and are row-normalized. On average, there are 23% closer and 39% less-close friends with $n = 100$. For the results reported below we consider two sets of network parameters as given in Table 1.

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

	Auroregressive Parameters							
	Equation 1				Equation 2			
	λ_{11}	λ_{12}	ρ_{11}	ρ_{12}	λ_{21}	λ_{22}	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

The remaining parameters are selected as $b_{12} = 0.3$, $b_{21} = 0.15$, $c_{11} = \ldots = c_{26} = 0.5$. The observations on the exogenous regressors are kept fixed for all MC iterations, and are generated as independent of each other and as cross sectionally *i.i.d.* N(1,3). The disturbances ε_1, ε_2 are generated as cross sectionally i.i.d. normal with mean 0, variance 1 and covariance .5.

In Table 2 we report the bias and root mean squared error (RMSE) of the LQ-GS2SLS and LQ-GS3SLS estimators and their refined counterparts, labled as Ref. LQ-GS2SLS and Ref. LQ-GS3SLS, for the two sets of parameters and $n = 100$. The results are based on $1,000$ MC runs each. To simplify the presentation we only report on the autoregressive network parameters of the two equations.

Table 2: Monte Carlo Simulation Results for Parameter Sets I and II and n=100

	LQ-GS2SLS		Ref. LQ-GS2SLS		LQ-GS3SLS		Ref. LQ-GS3SLS	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
				Parameter Set I				
Eq. 1:								
λ_{11}	-0.0100	0.1437	0.0439	0.1019	0.0016	0.1431	0.0566	0.1072
λ_{12}	-0.0881	0.1513	-0.1021	0.1386	-0.0756	0.1450	-0.0890	0.1286
ρ_{11}	-0.0295	0.2835	-0.1039	0.2284	-0.0257	0.2781	-0.1038	0.2285
ρ_{12}	0.0569	0.2633	0.0832	0.2382	0.0534	0.2652	0.0636	0.2275
Eq. 2:								
λ_{21}	0.0065	0.1012	0.0349	0.0964	0.0151	0.0973	0.0465	0.0936
λ_{22}	-0.0491	0.1070	-0.0731	0.1133	-0.0405	0.1065	-0.0612	0.1033
ρ_{21}	-0.0365	0.2355	-0.0838	0.2238	-0.0344	0.2274	-0.0865	0.2167
ρ_{22}	0.0178	0.2460	0.0553	0.2380	0.0169	0.2429	0.0371	0.2314
				Parameter Set II				
Eq. 1:								
λ_{11}	-0.0082	0.1247	0.0144	0.1096	-0.0145	0.1327	0.0104	0.1125
λ_{12}	-0.0105	0.0816	-0.0142	0.0738	-0.0130	0.0871	-0.0220	0.0774
ρ_{11}	-0.0021	0.2555	-0.0260	0.2238	0.0037	0.2521	-0.0224	0.2353
ρ_{12}	-0.0242	0.2710	-0.0028	0.2416	-0.0232	0.2708	-0.0033	0.2431
Eq. 2:								
λ_{21}	-0.0433	0.1260	-0.0061	0.1027	-0.0408	0.1292	-0.0026	0.0980
λ_{22}	-0.0273	0.0953	-0.0323	0.0764	-0.0235	0.1004	-0.0290	0.0809
ρ_{21}	0.0321	0.2615	-0.0197	0.2090	0.0281	0.2665	-0.0201	0.2188
ρ_{22}	-0.0111	0.2762	-0.0088	0.2384	-0.0132	0.2907	-0.0081	0.2525

As expected, in general, in terms of RMSE the full-information estimators dominate the respective limited-information estimators. What is of key interest here are the performances of LQ-GS2SLS and LQ-GS3SLS relative to their refined counterparts. The gains of the refined approach in terms of RMSE are quite substantial, particularly for Parameter Set II. Although for all estimators the biases are fairly small, we note that on average, relative to their refined counterparts, the biases of LQ-GS2SLS and LQ-GS3SLS are smaller. In an Online Appendix we report on additional results. One would expect that the bias and efficiency gains are smaller with stronger identification and smaller with larger sample sizes. We document this in Table A1 and A2 in the Online Appendix. Table A1 reports on $n = 100$, but in addition to the weaker identified case with $c_{gk} = 0.5$ the table also presents results for the stronger identified case with $c_{gk} = 1$. Table A2 reports results for both the weaker and stronger identified case, but for the larger sample size $n = 250$. The results support the conjecture that the use of the refined LQ-GS2SLS and LQ-GS3SLS can be especially beneficial for reducing the RMSE when identification is weak, but the benefits will mostly show up for relatively small sample sizes.

We think that the insights of this paper can also be relevant for other estimation procedures and classes of Anselin-Cliff-Ord-type models, where the explicit suppression of terms that converge to zero may help with improving the small sample performance of an estimation procedure.

References

[1] Anselin L (1988) Spatial econometrics: Methods and models. Kluwer Academic Publishers.

[2] Baltagi BH, Deng Y (2015) EC3SLS estimator for a simultaneous system of spatial autoregressive equations with random effects. Econometric Rev 34, 659-694.

[3] Cliff A, Ord J (1973) Spatial autocorrelation. Pion.

[4] Drukker DM, Egger PH, Prucha IR (2022) Simultaneous equations models with higher-order spatial or social network interactions. Econometric Theory, forthcoming.

[5] Kelejian HH, Prucha IR (2004) Estimation of simultaneous systems of spatially interrelated cross sectional equations. J Econometrics 118, 27-50.

[6] Liu X (2020) GMM identification and estimation of peer effects in a system of simultaneous equations. J Spatial Econometrics 1, 2-27.

[7] Yang K, Lee L-F (2019) Identification and estimation of spatial dynamic panel simultaneous equations models. Reg Sci Urb Econ 76, 32-46.

Online Appendix for Refined GMM Estimators for Simultaneous Equations Models with Network Interactions

Peter H. Egger[1], and Ingmar R. Prucha[2],

March 8, 2023

[1]ETH Zurich, Zurich, Switzerland; Tel.: +41-44-632-4108, e-mail: pegger@ethz.ch
[2]Department of Economics, University of Maryland, College Park, MD 20742, Tel.: +1-301-405-3499, e-mail: prucha@umd.edu

In this Online Appendix we report in Tables A1 and A2 on an extended set of Monte Carlo simulations. The simulations are intended to support some of the claims made in the main paper. In particular, we extend our study in the following ways:

- The results in Table 2 of the main paper are generated for $c_{gk} = 0.5$, where the c_{gk} denote the parameters on the exogenous variables. In this Online Appendix we also report results for $c_{gk} = 1$, noting that the latter parameter values imply stronger identification.
- The results in Table 2 of the main paper are generated for a sample size of $n = 100$. In this Online Appendix we also report results for $n = 250$.

The results in Tables A1 and A2 indicate that, indeed, the benefit of the proposed refinement of the LQ-GS2SLS and LQ-GS3SLS estimators tends to become more significant as identification becomes less strong. Moreover, a comparison of Table A2 for $n = 250$ with Table A1 for $n = 100$ demonstrates that the benefit of the proposed refinement is larger in smaller samples.

Table A1 - Monte Carlo Simulation Results for Parameter Sets I and II and n=100

	LQ-GS2SLS		Ref. LQ-GS2SLS		LQ-GS3SLS		Ref. LQ-GS3SLS	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
				Parameter Set I with c_{gk}=1				
Eq. 1:								
λ_{11}	0.0014	0.0528	0.0131	0.0493	0.0047	0.0502	0.0156	0.0462
λ_{12}	-0.0285	0.0639	-0.0308	0.0617	-0.0192	0.0580	-0.0226	0.0557
ρ_{11}	-0.0116	0.1820	-0.0423	0.1783	-0.0140	0.1814	-0.0387	0.1733
ρ_{12}	-0.0140	0.2280	0.0000	0.2068	-0.0163	0.2207	-0.0117	0.2030
Eq. 2:								
λ_{21}	0.0060	0.0398	0.0109	0.0397	0.0087	0.0376	0.0123	0.0380
λ_{22}	-0.0165	0.0461	-0.0206	0.0465	-0.0105	0.0411	-0.0133	0.0407
ρ_{21}	-0.0274	0.1678	-0.0406	0.1638	-0.0243	0.1664	-0.0379	0.1603
ρ_{22}	-0.0199	0.2112	-0.0126	0.2019	-0.0263	0.2089	-0.0160	0.1974
				Parameter Set I with c_{gk}=0.5				
Eq. 1:								
λ_{11}	-0.0100	0.1437	0.0439	0.1019	0.0016	0.1431	0.0566	0.1072
λ_{12}	-0.0881	0.1513	-0.1021	0.1386	-0.0756	0.1450	-0.0890	0.1286
ρ_{11}	-0.0295	0.2835	-0.1039	0.2284	-0.0257	0.2781	-0.1038	0.2285
ρ_{12}	0.0569	0.2633	0.0832	0.2382	0.0534	0.2652	0.0636	0.2275
Eq. 2:								
λ_{21}	0.0065	0.1012	0.0349	0.0964	0.0151	0.0973	0.0465	0.0936
λ_{22}	-0.0491	0.1070	-0.0731	0.1133	-0.0405	0.1065	-0.0612	0.1033
ρ_{21}	-0.0365	0.2355	-0.0838	0.2238	-0.0344	0.2274	-0.0865	0.2167
ρ_{22}	0.0178	0.2460	0.0553	0.2380	0.0169	0.2429	0.0371	0.2314
				Parameter Set II with c_{gk}=1				
Eq. 1:								
λ_{11}	0.0018	0.0696	0.0099	0.0649	0.0003	0.0616	0.0078	0.0555
λ_{12}	-0.0153	0.0596	-0.0172	0.0558	-0.0136	0.0558	-0.0145	0.0524
ρ_{11}	-0.0306	0.2331	-0.0493	0.2207	-0.0219	0.2277	-0.0404	0.2153
ρ_{12}	-0.0138	0.2629	-0.0129	0.2554	-0.0169	0.2598	-0.0112	0.2512
Eq. 2:								
λ_{21}	-0.0082	0.0711	0.0046	0.0594	-0.0082	0.0633	0.0036	0.0558
λ_{22}	-0.0097	0.0709	-0.0114	0.0605	-0.0080	0.0626	-0.0097	0.0541
ρ_{21}	-0.0030	0.2269	-0.0350	0.1992	-0.0040	0.2204	-0.0285	0.1881
ρ_{22}	-0.0181	0.2571	-0.0105	0.2271	-0.0214	0.2513	-0.0168	0.2277
				Parameter Set II with c_{gk}=0.5				
Eq. 1:								
λ_{11}	-0.0082	0.1247	0.0144	0.1096	-0.0145	0.1327	0.0104	0.1125
λ_{12}	-0.0105	0.0816	-0.0142	0.0738	-0.0130	0.0871	-0.0220	0.0774
ρ_{11}	-0.0021	0.2555	-0.0260	0.2238	0.0037	0.2521	-0.0224	0.2353
ρ_{12}	-0.0242	0.2710	-0.0028	0.2416	-0.0232	0.2708	-0.0033	0.2431
Eq. 2:								
λ_{21}	-0.0433	0.1260	-0.0061	0.1027	-0.0408	0.1292	-0.0026	0.0980
λ_{22}	-0.0273	0.0953	-0.0323	0.0764	-0.0235	0.1004	-0.0290	0.0809
ρ_{21}	0.0321	0.2615	-0.0197	0.2090	0.0281	0.2665	-0.0201	0.2188
ρ_{22}	-0.0111	0.2762	-0.0088	0.2384	-0.0132	0.2907	-0.0081	0.2525

Table A2 - Monte Carlo Simulation Results for Parameter Sets I and II and n=250

	LQ-GS2SLS		Ref. LQ-GS2SLS		LQ-GS3SLS		Ref. LQ-GS3SLS	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Parameter Set I with c_{gk}=1								
Eq. 1:								
λ_{11}	0.0002	0.0228	0.0008	0.0232	0.0030	0.0232	0.0032	0.0233
λ_{12}	-0.0053	0.0240	-0.0056	0.0240	-0.0034	0.0212	-0.0036	0.0210
ρ_{11}	-0.0071	0.0892	-0.0078	0.0890	-0.0065	0.0862	-0.0094	0.0873
ρ_{12}	0.0051	0.0973	0.0039	0.0974	0.0031	0.0985	0.0004	0.0983
Eq. 2:								
λ_{21}	0.0019	0.0220	0.0028	0.0220	0.0037	0.0203	0.0044	0.0207
λ_{22}	-0.0071	0.0229	-0.0073	0.0226	-0.0045	0.0214	-0.0048	0.0209
ρ_{21}	-0.0052	0.0988	-0.0080	0.0990	-0.0050	0.0999	-0.0084	0.0991
ρ_{22}	0.0043	0.1127	0.0028	0.1135	0.0015	0.1108	0.0009	0.1108
Parameter Set I with c_{gk}=0.5								
Eq. 1:								
λ_{11}	-0.0018	0.0461	0.0023	0.0471	0.0084	0.0480	0.0142	0.0485
λ_{12}	-0.0197	0.0511	-0.0229	0.0524	-0.0151	0.0468	-0.0167	0.0468
ρ_{11}	-0.0107	0.1040	-0.0151	0.1086	-0.0158	0.1058	-0.0230	0.1077
ρ_{12}	0.0193	0.1076	0.0204	0.1085	0.0138	0.1077	0.0152	0.1089
Eq. 2:								
λ_{21}	0.0081	0.0458	0.0115	0.0454	0.0125	0.0433	0.0181	0.0459
λ_{22}	-0.0229	0.0498	-0.0254	0.0497	-0.0156	0.0456	-0.0179	0.0441
ρ_{21}	-0.0205	0.1187	-0.0277	0.1194	-0.0221	0.1161	-0.0292	0.1176
ρ_{22}	0.0172	0.1324	0.0184	0.1315	0.0106	0.1271	0.0109	0.1261
Parameter Set II with c_{gk}=1								
Eq. 1:								
λ_{11}	0.0006	0.0294	0.0010	0.0297	0.0015	0.0287	0.0015	0.0279
λ_{12}	-0.0030	0.0319	-0.0031	0.0323	-0.0029	0.0300	-0.0028	0.0301
ρ_{11}	-0.0075	0.1107	-0.0094	0.1098	-0.0055	0.1037	-0.0074	0.1017
ρ_{12}	0.0063	0.1250	0.0060	0.1231	0.0023	0.1271	0.0013	0.1244
Eq. 2:								
λ_{21}	0.0008	0.0289	0.0018	0.0290	0.0011	0.0267	0.0017	0.0268
λ_{22}	-0.0044	0.0356	-0.0050	0.0353	-0.0036	0.0323	-0.0036	0.0323
ρ_{21}	-0.0048	0.1072	-0.0056	0.1077	-0.0017	0.1045	-0.0026	0.1054
ρ_{22}	0.0079	0.1349	0.0070	0.1342	0.0042	0.1293	0.0041	0.1286
Parameter Set II with c_{gk}=0.5								
Eq. 1:								
λ_{11}	0.0022	0.0630	0.0029	0.0644	0.0007	0.0610	0.0029	0.0606
λ_{12}	-0.0126	0.0667	-0.0125	0.0663	-0.0101	0.0638	-0.0092	0.0624
ρ_{11}	0.0003	0.1292	-0.0031	0.1282	0.0002	0.1254	-0.0045	0.1226
ρ_{12}	0.0159	0.1453	0.0158	0.1409	0.0151	0.1435	0.0147	0.1388
Eq. 2:								
λ_{21}	0.0034	0.0619	0.0076	0.0627	0.0007	0.0580	0.0038	0.0573
λ_{22}	-0.0147	0.0735	-0.0176	0.0740	-0.0111	0.0704	-0.0139	0.0692
ρ_{21}	-0.0039	0.1261	-0.0093	0.1251	-0.0049	0.1241	-0.0112	0.1217
ρ_{22}	0.0143	0.1610	0.0157	0.1580	0.0052	0.1604	0.0054	0.1542