Accepted Version for Regional Science and Urban Economics

Modified QML Estimation of Spatial Autoregressive Models with Unknown Heteroskedasticity and Nonnormality*

Shew Fan Liu and Zhenlin Yang†

School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903.

emails: shewfan.liu.2011@smu.edu.sg; zlyang@smu.edu.sg

January, 2015

Abstract

In the presence of heteroskedasticity, Lin and Lee (2010) show that the quasi maximum likelihood (QML) estimator of the spatial autoregressive (SAR) model can be inconsistent as a 'necessary' condition for consistency can be violated, and thus propose robust GMM estimators for the model. In this paper, we first show that this condition may hold in certain situations and when it does the regular QML estimator can still be consistent. In cases where this condition is violated, we propose a simple modified QML estimation method robust against unknown heteroskedasticity. In both cases, asymptotic distributions of the estimators are derived, and methods for estimating robust variances are given, leading to robust inferences for the model. Extensive Monte Carlo results show that the modified QML estimator outperforms the GMM and QML estimators even when the latter is consistent. The proposed methods are then extended to the more general SARAR models.

Key Words: Spatial dependence; Unknown heteroskedasticity; Nonnormality, Modified QML estimator; Robust standard error; SARAR models.

JEL Classification: C10, C13, C15, C21

1. Introduction

Spatial dependence is increasingly becoming an integral part of empirical works in economics as a means of modelling the effects of 'neighbours' (see, e.g., Cliff and Ord (1972, 1973, 1981), Ord (1975), Anselin (1988, 2003), Anselin and Bera (1998), LeSage and Pace (2009) for some early and comprehensive works). Spatial interaction in general can occur in many forms. For instance peer interaction can cause stratified behaviour in the sample such as herd behaviour in stock markets, innovation spillover effects, localized purchase decisions, etc., while spatial relationships can also occur more naturally due to structural differences in space/cross-section such as geographic proximity, trade agreements, demographic characteristics, etc. See Case (1991), Pinkse and Slade (1998), Pinkse et al. (2002), Hanushek et al. (2003), Baltagi et al. (2007) to name a few. Among the various spatial econometrics models that have been extensively treated, the most popular one may be the spatial autoregressive (SAR) model.

*The authors are grateful to School of Economics, Singapore Management University, for research support, and to Editor, Daniel McMillen, a referee, and the participants of the VIII world conference of the Spatial Econometrics Association, Zürich 2014, for helpful comments.

†Corresponding author: 90 Stamford Road, Singapore 178903. Phone +65-6828-0852; Fax: +65-6828-0833.

While heteroskedasticity is common in regular cross-section studies, it may be more so for a spatial econometrics model due to aggregation, clustering, etc. Anselin (1988) identifies that heteroskedasticity can broadly occur due to "idiosyncrasies in model specification and affect the statistical validity of the estimated model". This may be due to the misspecification of the model that feeds to the disturbance term or may occur more naturally in the presence of peer interactions. Data related heteroskedasticity may also occur for example if the model deals with a mix of aggregate and non aggregate data, the aggregation may cause errors to be heteroskedastic. See, e.g., Glaeser et al. (1996), LeSage and Pace (2009), Lin and Lee (2010), Kelejian and Prucha (2010), for more discussions. As such, the assumption of homoskedastic disturbances is likely to be invalid in a spatial context in general. However, much of the present spatial econometrics literature has focused on estimators developed under the assumption that the errors are homoskedastic. This is in a clear contrast to the standard cross-section econometrics literature where the use of heteroskedasticity robust estimators is a standard practice.

Although Anselin raised the issue of heteroskedasticity in spatial models as early as in 1988, and made an attempt to provide tests of spatial effects robust to unknown heteroskedasticity, comprehensive treatments of estimation related issues were not considered until recent years by, e.g., Kelejian and Prucha (2007, 2010), LeSage (1997), Lin and Lee (2010), Arraiz et al. (2010), Badinger and Egger (2011), Jin and Lee (2012), Baltagi and Yang (2013b), and Doğan and Taşpinar (2014). Lin and Lee (2010) formally illustrate that the traditional quasi maximum likelihood (QML) and generalized method of moments (GMM) estimators are inconsistent in general when the SAR model suffers from heteroskedasticity, and provide heteroskedasticity robust GMM estimators by modifying the usual quadratic moment conditions.

Inspired by Lin and Lee (2010), we introduce a modified QML estimator (QMLE) for the SAR model by modifying the concentrated score function for the spatial parameter to make it robust against unknown heteroskedasticity. It turns out that the method is very simple and more importantly, it can be easily generalized to suit more general models.[1] For heteroskedasticity robust inferences, we propose an outer-product-of-gradient (OPG) method for estimating the variance of the modified QMLE. We provide formal theories for the consistency and asymptotic normality of the proposed estimator, and the consistency of the robust standard error estimate. Extensive Monte Carlo results show that the modified QML estimator generally outperforms its GMM counter parts in terms of efficiency and sensitivity to the magnitude of model parameters in particular the regression coefficients. The Monte Carlo results also show that the proposed robust standard error estimate performs well. We also study the cases under which the regular QMLE is robust against unknown heteroskedasticity and provide a set of robust inference methods. It is interesting to note that the modified QMLE is computationally as simple as the regular QMLE, and it also outperforms the regular QMLE when the latter is heteroskedasticity robust. This is because the modified QMLE captures the extra variability inherent from the

[1] The efficiency of an MLE may be the driving force for exploiting a likelihood-based estimator for achieving robustness against various model misspecifications such as heteroskedasticity and nonnormality. The computational complexity may be the key factor that hinders the application of the ML-type estimation method. However, with the modern computing technologies this is no longer an issue of major concern, unless n is very large.

estimation of the regression coefficients and the average of error variances.

To demonstrate their flexibility and generality, the proposed methods are then extended to the popular spatial autoregressive model with spatial autoregressive disturbances (SARAR(1,1)) with heteroskedastic innovations. Kelejian and Prucha (2010) formally treat this model with a three-step estimation procedure. Monte Carlo results show that the modified QMLE performs better in finite sample than the three-step estimator. Further possible extensions of the proposed methods are discussed. In summary, the proposed set of QML-based robust inference methods are simple and reliable, and can be easily adopted by applied researchers.

The rest of the paper is organized as follows. Section 2 examines the cases where the regular QML estimator of the SAR model is consistent under unknown heteroskedasticity, and provides methods for robust inferences. Section 3 introduces the modified QML estimator that is generally robust against unknown heteroskedasticity, and presents methods for robust inferences. Section 4 presents the Monte Carlo results for the SAR model. Section 5 extends the proposed methods to the popular SARAR(1,1) model and discusses further possible extensions. Section 6 concludes the paper. All technical details are given in Appendix B.

2. QML Estimation of Spatial Autoregressive Models

In this section, we first outline the QML estimation of the SAR model under the assumptions that the errors are independent and identically distributed (iid). Then, we examine the properties of the QMLE of the SAR model when the errors are independent but not identically distributed (inid). We provide conditions under which the regular QMLE is robust against heteroskedasticity of unknown form, derive its asymptotic distribution, and provide heteroskedasticity robust estimator of its asymptotic variance.

Some general notation will be followed in this paper: $|\cdot|$ and $\text{tr}(\cdot)$ denote, respectively, the determinant and trace of a square matrix; A' denotes the transpose of a matrix A; $\text{diag}(\cdot)$ denotes the diagonal matrix formed by a vector or the diagonal elements of a square matrix; $\text{diagv}(\cdot)$ denotes the column vector formed by the diagonal elements of a square matrix; and a vector raised to a certain power is operated elementwise.

2.1 The model and the QML estimation

Consider the spatial autoregressive or SAR model of the form:

$$Y_n \;=\; \lambda_0 W_n Y_n + X_n\beta_0 + \epsilon_n, \tag{1}$$

where X_n is an $n\times k$ matrix of exogenous variables, W_n is a known $n\times n$ spatial weights matrix, ϵ_n is an $n\times 1$ vector of disturbances of independent and identically distributed (iid) elements with mean zero and variance σ^2, β is a $k\times 1$ vector of regression coefficients and λ is the spatial parameter. The Gaussian loglikelihood of $\theta=(\beta',\sigma^2,\lambda)'$ is,

$$\ell_n(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) + \ln|A_n(\lambda)| - \frac{1}{2\sigma^2}\epsilon_n'(\beta,\lambda)\epsilon_n(\beta,\lambda), \tag{2}$$

where $A_n(\lambda) = I_n - \lambda W_n$, I_n is an $n \times n$ identity matrix, and $\epsilon_n(\beta,\lambda) = A_n(\lambda)Y_n - X_n\beta$. Given λ, $\ell_n(\theta)$ is maximized at $\hat{\beta}_n(\lambda) = (X_n'X_n)^{-1}X_n'A_n(\lambda)Y_n$ and $\hat{\sigma}_n^2(\lambda) = \frac{1}{n}Y_n'A_n'(\lambda)M_nA_n(\lambda)Y_n$, where $M_n = I_n - X_n(X_n'X_n)^{-1}X_n'$. By substituting $\hat{\beta}_n(\lambda)$ and $\hat{\sigma}_n^2(\lambda)$ into $\ell_n(\theta)$, we arrive at the concentrated Gaussian loglikelihood function for λ as,

$$\ell_n^c(\lambda) = -\frac{n}{2}[\ln(2\pi)+1] - \frac{n}{2}\ln(\hat{\sigma}_n^2(\lambda)) + \ln|A_n(\lambda)|. \tag{3}$$

Maximizing $\ell_n^c(\lambda)$ gives the unconstrained QMLE $\hat{\lambda}_n$ of λ, and thus the QMLEs of β and σ^2 as $\hat{\beta}_n \equiv \hat{\beta}(\hat{\lambda}_n)$ and $\hat{\sigma}_n^2 \equiv \hat{\sigma}_n^2(\hat{\lambda}_n)$. Denote $\hat{\theta}_n = (\hat{\beta}_n', \hat{\sigma}_n^2, \hat{\lambda}_n)'$, the QMLE of θ.

Under regularity conditions, Lee (2004) establishes the consistency and asymptotic normality of the QMLE $\hat{\theta}_n$. In particular, he shows that $\hat{\lambda}_n$ and $\hat{\beta}_n$ may have a slower than $\sqrt{n}$-rate of convergence if the degree of spatial dependence (or the number of neighbours each spatial unit has) grows with the sample size n. The QMLE and its asymptotic distribution developed by Lee are robust against nonnormality of the error distribution. However, some important issues need to be further considered: (i) conditions under which the regular QMLE $\hat{\theta}_n$ remains consistent when errors are heteroskedastic, (ii) methods to modify the regular QMLE $\hat{\theta}_n$ so that it becomes generally consistent under unknown heteroskedasticity, and (iii) methods for estimating the variance of the (modified) QMLE robust against unknown heteroskedasticity.

2.2 Robustness of QMLE against unknown heteroskedasticity

It is accepted that the regular QMLE of the usual linear regression model without spatial dependence, developed under homoskedastic errors, is still consistent when the errors are in fact heteroskedastic. However, for correct inferences the standard error of the estimator has to be adjusted to account for this unknown heteroskedasticity (White, 1980). Suppose now we have a linear regression model with spatial dependence as given in (1) with disturbances that are inid with means zero and variances $\sigma^2 h_{n,i}$, $i = 1,\ldots,n$, where $h_{n,i} > 0$ and $\frac{1}{n}\sum_{i=1}^n h_{n,i} = 1$.[2] Consider the score function derived from (2),

$$\psi_n(\theta) = \frac{\partial \ell_n(\theta)}{\partial\theta} = \begin{cases} \frac{1}{\sigma^2}X_n'\epsilon_n(\beta,\lambda), \\ \frac{1}{2\sigma^4}[\epsilon_n'(\beta,\lambda)\epsilon_n(\beta,\lambda) - n\sigma^2], \\ \frac{1}{\sigma^2}Y_n'W_n'\epsilon_n(\beta,\lambda) - \mathrm{tr}[G_n(\lambda)], \end{cases} \tag{4}$$

where $G_n(\lambda) = W_nA^{-1}(\lambda)$. It is well known that for an extremum estimator, such as the QMLE $\hat{\theta}_n$ we consider, to be consistent, a necessary condition is that $\mathrm{plim}_{n\to\infty}\frac{1}{n}\psi_n(\theta_0) = 0$ at the true parameter θ_0 (Amemiya, 1985). This is always the case for the β and σ^2 components of $\psi_n(\theta_0)$ whether or not the errors are homoskedastic. However, it may not be the case for the λ component of $\psi_n(\theta_0)$. Let $h_n = (h_{n,1},\ldots,h_{n,n})'$, $g_n = (g_{n,1},\ldots,g_{n,n})' = \mathrm{diagv}(G_n)$, $\bar{g}_n = \frac{1}{n}\sum_{i=1}^n g_{n,i}$, $H_n = \mathrm{diag}(h_n)$. Let $\mathrm{Cov}(g_n, h_n)$ denote the sample covariance between the

[2] Note that σ^2 is the average of $\mathrm{Var}(\epsilon_{n,i})$. Under homoskedasticity, $h_{n,i} = 1, \forall i$. For generality, we allow $h_{n,i}$ to depend on n, for each i. This parameterization, a nonparametric version of Breusch and Pagan (1979), is useful as it allows the estimation of the average scale parameter. See Section 3 for more details.

two vectors g_n and h_n. We have, similarly to Lin and Lee (2010),

$$\begin{aligned} \tfrac{1}{n}\tfrac{\partial}{\partial\lambda}\ell_n(\theta_0) &= \tfrac{1}{n}\mathrm{tr}(H_nG_n - G_n) + o_p(1) \\ &= \tfrac{1}{n}\textstyle\sum_{i=1}^n (h_{n,i}-1)(g_{n,i}-\bar{g}_n) + o_p(1) \\ &= \mathrm{Cov}(g_n, h_n) + o_p(1). \end{aligned} \tag{5}$$

Therefore, for $\hat{\theta}_n$ to be consistent, it is necessary that as $n \to \infty$, $\mathrm{Cov}(g_n, h_n) \to 0$; in other words, when $\lim_{n\to\infty} \mathrm{Cov}(g_n, h_n) \neq 0$, $\hat{\theta}_n$ cannot be consistent.

Lin and Lee (2010) noted that this condition is satisfied if almost all the diagonal elements of the matrix G_n are equal. In fact, by Cauchy-Schwartz inequality, this condition is satisfied if $\mathrm{Var}(g_n) \to 0$, which boils down to $\mathrm{Var}(k_n) \to 0$, where k_n is the vector of number of neighbours for each unit.[3] Furthermore, if heteroskedasticity occurs due to reasons unrelated to the number of neighbours, for example, due to the nature of the exogenous regressors X_n, then the required condition will still be satisfied. These discussions suggest that the regular QMLE of the SAR model derived under homoskedasticity can still be consistent when in fact the errors are heteroskedastic. Formal results in this context can be constructed under the following regularity conditions. A quantity defined at the true parameter is represented with a suppressed variable notation, e.g., $A_n \equiv A_n(\lambda_0)$ and $G_n \equiv G_n(\lambda_0)$.

Assumption 1: *The true parameter λ_0 is in the interior of a compact parameter set Λ.*[4]

Assumption 2: *$\epsilon_n \sim (0, \sigma_0^2 H_n)$, where $H_n = \mathrm{diag}(h_{n,1}, \ldots, h_{n,n})$, such that $\frac{1}{n}\sum_{i=1}^n h_{n,i} = 1$ and $h_{n,i} > 0, \forall i$ and $E|\epsilon_{n,i}|^{4+\delta} < c$ for some $\delta > 0$ and constant c for all n and i.*

Assumption 3: *The elements of the $n \times k$ regressor matrix X_n are uniformly bounded for all n, X_n has the full rank k, and $\lim_{n\to\infty} \frac{1}{n}X_n'X_n$ exists and is nonsingular.*

Assumption 4: *The spatial weights matrix W_n is uniformly bounded in absolute value in both row and column sums and its diagonal elements are zero.*

Assumption 5: *The matrix A_n is non-singular and A_n^{-1} is uniformly bounded in absolute value in both row and column sums. Further, $A_n^{-1}(\lambda)$ is uniformly bounded in either row or column sums, uniformly in $\lambda \in \Lambda$.*

Assumption 6: *The limit $\lim_{n\to\infty} \frac{1}{n}(G_nX_n\beta_0)'M_n(G_nX_n\beta_0) = k$, where either $k > 0$, or $k = 0$ but $\lim_{n\to\infty} \frac{1}{n}\ln|\sigma_0^2 A_n^{-1}A_n'^{-1}| - \frac{1}{n}\ln|\sigma_n^2(\lambda)A_n^{-1}(\lambda)A_n'^{-1}(\lambda)| \neq 0$, whenever $\lambda \neq \lambda_0$, where $\sigma_n^2(\lambda) = \frac{1}{n}\sigma_0^2 \mathrm{tr}(H_nA_n'^{-1}A_n'^{-1}(\lambda)A_n^{-1}(\lambda)A_n^{-1})$.*

[3] This is because (*i*) $G_n = W_n + \lambda W_n^2 + \lambda^2 W_n^3 + \ldots$, if $|\lambda| < 1$ and $w_{n,ij} < 1$, and (*ii*) the diagonal elements of $W_n^r, r \geq 2$ inversely relate to k_n, see Anselin (2003). In fact, when W_n is row-normalized and symmetric, $\mathrm{diag}(W_n^2) = \{k_{n,i}^{-1}\}$. $\mathrm{Var}(k_n) = o(1)$ can be seen to be true for many popular spatial layouts such as Rook, Queen, group interactions such that variation in groups sizes becomes small when n gets large, etc, see Yang (2010).

[4] For QML-type estimation, the parameter space Λ must be such that $A_n(\lambda)$ is non-singular $\forall \lambda \in \Lambda$. If the eigenvalues of W_n are all real, then $\Lambda = (w_{\min}^{-1}, w_{\max}^{-1})$ where $w_{\min}$ and $w_{\max}$ are, respectively, the smallest and the largest eigenvalues of W_n; if, W_n is row normalized, then $w_{\max} = 1$ and $w_{\min}^{-1} < -1$, and $\Lambda = (w_{\min}^{-1}, 1)$ (Anselin, 1988). In general, the eigenvalues of W_n may not be all real as W_n can be asymmetric. LeSage and Pace (2009, p. 88-89) argue that only the purely real eigenvalues can affect the singularity of $A_n(\lambda)$. Consequently, for W_n with complex eigenvalues, the interval of λ that guarantees non-singular $A_n(\lambda)$ is $(w_s^{-1}, 1)$ where w_s is the most negative real eigenvalue of W_n. Kelejian and Prucha (2010) suggest Λ be $(-\tau_n^{-1}, \tau_n^{-1})$ where τ_n is the spectral radius of W_n, or $(-1, 1)$ after normalization.

Assumptions 2 and 3 are similar to those from Lin and Lee (2010). Assumption 2 implies that $\{h_{n,i}\}$ as well as the third and fourth moments of $\epsilon_{n,i}$ are uniformly bounded for all n and i. Assumptions 2 and 3 imply that $\lim_{n\to\infty}\frac{1}{n}X_n'H_nX_n$ exists and is nonsingular. Assumptions 4 and 5 are standard for the SAR model. The uniform boundedness conditions limit the spatial dependence to a manageable level (Kelejian and Prucha, 1999). Assumption 6 is the heteroskedastic version of the identification condition introduced by Lee (2004) for the homoskedastic SAR model.

For the loglikelihood and score functions given in (2) and (4), let $\mathbb{I}_n = -\frac{1}{n}\mathrm{E}\big[\frac{\partial^2}{\partial\theta\partial\theta'}\ell_n(\theta_0)\big]$ and $\Sigma_n = \frac{1}{n}\mathrm{E}\big[\frac{\partial}{\partial\theta}\ell_n(\theta_0)\frac{\partial}{\partial\theta'}\ell_n(\theta_0)\big]$, with their exact expressions deferred to the next subsection in connection with the issue on the robust variance covariance matrix estimation. We have the following results (recall $g_n = \mathrm{diagv}(G_n)$ and let $q_n = \mathrm{diagv}(G_n'G_n)$).

Theorem 1: *Under Assumptions 1-6, $Cov(g_n, h_n) = o(1)$ and $Cov(q_n, h_n) = o(1)$, we have as $n\to\infty$, $\hat\theta_n \xrightarrow{p} \theta_0$; under Assumptions 1-6 and $Cov(g_n, h_n) = o(n^{-1/2})$, we have as $n\to\infty$,*

$$\sqrt{n}(\hat\theta_n - \theta_0) \xrightarrow{D} N(0, \mathbb{I}^{-1}\Sigma\ \mathbb{I}^{-1}), \tag{6}$$

where $\mathbb{I} = \lim_{n\to\infty}\mathbb{I}_n$ and $\Sigma = \lim_{n\to\infty}\Sigma_n$ both assumed to exist and $\mathbb{I}$ is nonsingular.

2.3 Robust standard errors of the QML estimators

Asymptotically valid inference for θ based on the QMLEs $\hat\theta_n$ requires a consistent estimator of the asymptotic variance given in Theorem 1. This is fairly simple under homoskedasticity as the sample analogue of $\mathbb{I}_n$ and Σ_n can directly be used to give consistent estimators of $\mathbb{I}$ and Σ. Under the unknown heteroskedasticity designated by H_n, we have after some algebra:

$$\mathbb{I}_n = \begin{pmatrix} \frac{1}{\sigma_0^2 n}X_n'X_n & 0 & \frac{1}{\sigma_0^2 n}X_n'\eta_n \\ \sim & \frac{1}{2\sigma_0^4} & \frac{1}{\sigma_0^2 n}\mathrm{tr}(H_nG_n) \\ \sim & \sim & \frac{1}{\sigma_0^2 n}\eta_n'\eta_n + \frac{1}{n}\mathrm{tr}(H_nG_n'G_n + G_n^2) \end{pmatrix},$$

where $\eta_n = G_nX_n\beta_0$. This shows that a consistent estimator of $\mathbb{I}_n$ can still be obtained by 'plugging' $\hat\theta_n$ for θ_0, $G_n(\hat\theta_n)$ for G_n and $\hat H_n = \frac{1}{\hat\sigma_n^2}\mathrm{diag}(\hat\epsilon_{n,1}^2, \dots, \hat\epsilon_{n,n}^2)$ for H_n, in line with the idea of White (1980), where $\{\hat\epsilon_{n,i}\}$ are the QML residuals. However, this approach fails in estimating the variance of the score, Σ_n, as its σ_0^2-element:

$$\Sigma_{n,\sigma^2\sigma^2} = \frac{1}{4n\sigma_0^4}\sum_{i=1}^n(\kappa_{n,i} + 2h_{n,i}^2),$$

cannot be consistently estimated unless the excess kurtosis measures $\{\kappa_{n,i}\}$ are all zero or $\{\epsilon_{n,i}\}$ are normally distributed. This means that the robust inference method for σ_0^2 is not available. Obviously, σ^2 is typically not the main parameter that inferences concern, although the consistency of its QMLE (shown in Theorem 1) is crucial. Thus, to get around this problem, we focus on λ and β as those are the main parameters that inferences concern. First, based on the concentrated score function for λ, obtained from (4) by concentrating out β and σ^2 (see

(7) below), we obtain the robust variance of $\hat{\lambda}_n$, and then based on the relationship between $\hat{\beta}_n$ and $\hat{\lambda}_n$ we obtain the robust variance of $\hat{\beta}_n$. As these developments fall into the main results presented in next section, we give details at the end of Section 3.3.

3. Modified QML Estimation under Heteroskedasticity

As argued in Lin and Lee (2010) and further discussed in Section 2 of this paper, the necessary condition for the consistency of the regular QMLE, $\lim_{n\to\infty} \text{Cov}(g_n, h_n) = 0$, can be violated when h_n is proportional to the number of neighbours k_n for each spatial unit and $\lim_{n\to\infty} \text{Var}(k_n) \neq 0$.[5] To solve this problem, Lin and Lee (2010) propose robust GMM and optimal robust GMM estimators for the SAR model. In this paper, we introduce a modified QMLE for the SAR model by modifying the concentrated score function for the spatial parameter to make it robust against unknown heteroskedasticity. It turns out that the method is very simple and more importantly it can be easily generalized to suit more general models (see Section 5). Furthermore, the method of modification takes into account the estimation of the β and σ^2 parameters, thus can be expected to have a good finite sample performance. Indeed, the Monte Carlo results presented in Section 4 show an excellent finite sample performance of the proposed estimator. For robust inferences concerning the spatial or regression parameters, we introduce OPG estimators of the variances of the modified QMLEs.

3.1 The modified QML Estimator

Given the problems associated with the λ-element of $\psi_n(\theta_0)$ in (4), in asymptotically attaining the limit desired to ensure consistency of the related extremum estimator under heteroskedasticity, one can look at a modification to the score function that allows it to reach a probability limit of zero by brute force. This method is in line with Lin and Lee (2010)'s treatment to the quadratic moments of the form $\text{E}(\epsilon_n' P_n \epsilon_n) = 0$, where $\text{tr}(P_n) = 0$ is modified such that $\text{diag}(P_n) = 0$ to attain a consistent GMM estimator under unknown heteroskedasticity. Following this idea, if we modify the last component of $\psi_n(\theta_0)$ as,

$$\sigma_0^{-2}[Y_n' W_n' \epsilon_n - \epsilon_n' \text{diag}(G_n) \epsilon_n],$$

we immediately see that $\text{plim} \frac{1}{n\sigma_0^2}[Y_n' W_n' \epsilon_n - \epsilon_n' \text{diag}(G_n) \epsilon_n] = 0$, in light of (5). This modification is asymptotically valid in the sense that it will make the estimators consistent under the unknown heteroskedasticity. However, the finite sample performance of the estimators is not guaranteed as the variations from the estimation of β and σ^2 are completely ignored.

Now consider the average concentrated score function derived by concentrating out β and σ^2, i.e., replacing β and σ^2 by $\hat{\beta}_n(\lambda)$ and $\hat{\sigma}_n^2(\lambda)$ in the last component of (4), or taking the

[5]For example, when W_n corresponds to group interactions (circular world spatial layout can be a special case), and the group sizes are generated from a fixed discrete distribution, we have $\lim_{n\to\infty} \text{Var}(k_n) \neq 0$. In fact, in many empirical situations, the spatial weight matrix is constructed from economic or geographic distance, and hence does not satisfy the condition $\text{Cov}(g_n, h_n) = o(1)$.

derivative of (3), and then dividing the resulting concentrated score function by n,

$$\tilde{\psi}_n(\lambda) = \frac{Y_n' A_n'(\lambda) M_n [G_n(\lambda) - \frac{1}{n}\mathrm{tr}(G_n(\lambda)) I_n] A_n(\lambda) Y_n}{Y_n' A_n'(\lambda) M_n A_n(\lambda) Y_n}. \tag{7}$$

The average concentrated score $\tilde{\psi}_n(\lambda)$ captures the variability coming from estimating β and σ^2. Under the regular QML estimation framework (see, e.g., Amemiya, 1985), the QMLE of λ is equivalently defined as $\hat{\lambda}_n = \arg\{\tilde{\psi}_n(\lambda) = 0\}$. Solving $\tilde{\psi}_n(\lambda) = 0$ is equivalent to solving $Y_n' A_n'(\lambda) M_n [G_n(\lambda) - \frac{1}{n}\mathrm{tr}(G_n(\lambda)) I_n] A_n(\lambda) Y_n = 0$, and for the solution $\hat{\lambda}_n$ to remain consistent under unknown heteroskedasticity, it is necessary that $\frac{1}{n}\mathrm{E}[Y_n' A_n' M_n (G_n - \frac{1}{n}\mathrm{tr}(G_n) I_n) A_n Y_n]$ equals or tends to zero, see van der Vaart (1998, ch. 5). This is not true if there exists unknown heteroskedasticity and the conditions stated in Theorem 1 are violated.

Our idea is to modify the numerator of (7) so that its expectation at the true parameter λ_0 is zero even under unknown heteroskedasticity.[6] Since $\mathrm{E}(Y_n' A_n' M_n G_n A_n Y_n) = \sigma_0^2 \mathrm{tr}(H_n M_n G_n) = \sigma_0^2 \mathrm{tr}(H_n \mathrm{diag}(M_n G_n))$, this suggests that one should replace $\frac{1}{n}\mathrm{tr}(G_n) I_n$ in the numerator of (7) by $\mathrm{diag}(M_n G_n)$. However, $\mathrm{E}(Y_n' A_n' M_n \mathrm{diag}(M_n G_n) A_n Y_n) = \sigma_0^2 \mathrm{tr}(H_n M_n \mathrm{diag}(M_n G_n)) \neq \mathrm{E}(Y_n' A_n' M_n G_n A_n Y_n)$. Thus, in order to cancel the effect of the additional M_n, one should instead replace $\frac{1}{n}\mathrm{tr}(G_n) I_n$ in the numerator of (7) by $\mathrm{diag}(M_n)^{-1} \mathrm{diag}(M_n G_n)$. Hence, $\tilde{\psi}_n(\lambda)$ is modified by replacing $G_n(\lambda) - \frac{1}{n}\mathrm{tr}(G_n(\lambda)) I_n$ by,

$$G_n^{\circ}(\lambda) = G_n(\lambda) - \mathrm{diag}(M_n)^{-1} \mathrm{diag}(M_n G_n(\lambda)). \tag{8}$$

This gives a modified concentrated score function,

$$\tilde{\psi}_n^*(\lambda) = \frac{Y_n' A_n'(\lambda) M_n G_n^{\circ}(\lambda) A_n(\lambda) Y_n}{Y_n' A_n'(\lambda) M_n A_n(\lambda) Y_n}, \tag{9}$$

and hence a modified QML estimator of λ_0 as,

$$\tilde{\lambda}_n = \arg\{\tilde{\psi}_n^*(\lambda) = 0\}. \tag{10}$$

Once a heteroskedasticity robust estimator of λ is obtained, the heteroskedasticity robust estimators (or the modified QMLEs) of β and σ^2 are, respectively, $\tilde{\beta}_n = \hat{\beta}_n(\tilde{\lambda}_n)$ and $\tilde{\sigma}_n^2 = \hat{\sigma}_n^2(\tilde{\lambda}_n)$ as the estimating functions (first two components of $\psi_n(\theta)$) leading to $\hat{\beta}_n(\lambda)$ and $\hat{\sigma}_n^2(\lambda)$ defined below (2) are robust to unknown heteroskedasticity. More discussions on this will follow.

Recently, Jin and Lee (2012) proposed a heteroskedasticity robust *root estimator* of λ by solving the quadratic (in λ) equation: $Y_n' A_n'(\lambda) M_n P_n A_n(\lambda) Y_n = 0$, where P_n is an $n \times n$ matrix such that $M_n P_n$ has a zero diagonal. As there are two roots and only one is consistent, they gave criteria to choose the consistent root. In case where the P_n matrix is parameter dependent, they suggested using some initial consistent estimates to come up with an estimate, say $\widehat{P}_n$, of

[6]Making the expectation of an estimating function to be zero leads potentially to a finite sample bias corrected estimation. This is in line with Baltagi and Yang (2013a,b) in constructing standardized or heteroskedasticity-robust LM tests with finite sample improvements. See also Kelejian and Prucha (2001, 2010) and Lin and Lee (2010) for some useful methods in handling the linear-quadratic forms of heteroskedastic random vectors.

P_n, and then solve $Y_n' A_n'(\lambda) M_n \widehat{P}_n A_n(\lambda) Y_n = 0$. Clearly, $G_n^{\circ}(\lambda)$ defined above is a choice for P_n although an initial estimate of λ, say $\hat{\lambda}_n^0$, is needed to obtain $\widehat{P}_n = G_n^{\circ}(\hat{\lambda}_n^0)$. Jin and Lee also suggest this. This approach is attractive as the root estimator has a closed-form expression and thus can handle a super large data. However, it can be ambiguous in practice in choosing a consistent root as the selection criterion is parameter dependent. Furthermore, our Monte Carlo simulation shows that $Y_n' A_n'(\lambda) M_n \widehat{P}_n A_n(\lambda) Y_n = 0$ tends to give non-real roots when $|\lambda|$ is not small, say ≥ 0.5, in particular when λ is negative, and when n is not very large. In contrast, this problem does not occur to the modified QML estimator $\tilde{\lambda}_n$ given above. Thus, the modified QML estimator $\tilde{\lambda}_n$ proposed in this paper complements Jin and Lee's (2012) root estimator. More discussions along this line are given in the following sections. Some remarks follow before moving into the asymptotic properties of the modified QML estimators.

Remark 1: It turns out that the modified QMLEs of the SAR model are computationally as simple as the original QMLEs, but the former are generally consistent under unknown heteroskedasticity while preserving the nature of being robust against nonnormality.

Remark 2: The proposed methods can be easily extended to more advanced models (spatial or non-spatial) as demonstrated in Section 5. However, it is not clear to us how to extend the GMM estimators of Lin and Lee (2010) to a more general model, and the root estimator of Jin and Lee (2012) may run into difficulty for a more general model as when there are two (or more) quadratic functions of two (or more) unknowns, it is difficult to choose the consistent roots.

Remark 3: The correction $G_n^{\circ}(\lambda) = G_n(\lambda) - \text{diag}(M_n)^{-1}\text{diag}(M_n G_n(\lambda))$ as opposed to the more intuitively appealing correction $G_n(\lambda) - \text{diag}(G_n(\lambda))$ has better finite sample performance since the modification is made directly on the concentrated score function which contains the variability accruing from the estimation of β and σ^2.

3.2 Asymptotic distribution of the modified QML estimators

To ensure that the modified estimation function given in (9) uniquely identifies λ_0, the Assumption 6 needs to be modified as follows. Let $\Omega_n(\lambda) = A_n'(\lambda)[G_n(\lambda) - \text{diag}(G_n(\lambda))]A_n(\lambda)$.

Assumption 6*: $\lim_{n\to\infty} \frac{1}{n}[\beta_0' X_n' A_n'^{-1} \Omega_n(\lambda) A_n^{-1} X_n \beta_0 + \sigma_0^2 \text{tr}(H_n A_n'^{-1} \Omega_n(\lambda) A_n^{-1})] \neq 0, \forall \lambda \neq \lambda_0$.

The central limit theorem (CLT) for linear quadratic forms of Kelejian and Prucha (2001) allows for heteroskedasticity and can be used to prove the asymptotic normality of the modified QML estimator. First, it is easy to show that the normalized and modified concentrated score function has the following representation at λ_0,

$$\sqrt{n}\tilde{\psi}_n^* \equiv \sqrt{n}\tilde{\psi}_n^*(\lambda_0) = \frac{1}{\sqrt{n}\sigma_0^2}\left(\epsilon_n' B_n \epsilon_n + c_n' \epsilon_n\right) + o_p(1), \tag{11}$$

where $B_n = M_n G_n^{\circ}$ and $c_n = M_n G_n^{\circ} X_n \beta_0$, because $\hat{\sigma}_n^2(\lambda_0) = \frac{1}{n}\epsilon_n' M_n \epsilon_n = \frac{1}{n}\text{E}(\epsilon_n' M_n \epsilon_n) + o_p(1) = \frac{\sigma_0^2}{n}\text{tr}(H_n M_n) + o_p(1) = \sigma_0^2 + o_p(1)$, and it follows that $\hat{\sigma}_n^{-2}(\lambda_0) = \sigma_0^{-2} + o_p(1)$.

Let $\tau_n(\cdot)$ denote the first-order standard deviation and $\tau_n^2(\cdot)$ the first-order variance of a normalized quantity, e.g., $\tau_n^2(\tilde{\psi}_n^*)$ is the first-order term of $\text{Var}(\sqrt{n}\tilde{\psi}_n^*)$, and $\tau_n^2(\tilde{\lambda}_n)$ is the first-order term of $\mathrm{Var}(\sqrt{n}\tilde{\lambda}_n)$. By the representation (11) and Lemma A.3, we have,

$$\tau_n^2(\tilde{\psi}_n^*) = \tfrac{1}{n}\textstyle\sum_{i=1}^n (b_{n,ii}^2 h_{n,i}^2 \kappa_{n,i} + \tfrac{2}{\sigma_0^4} b_{n,ii} c_{n,i} s_{n,i}) + \tfrac{1}{n}\mathrm{tr}[H_n B_n (H_n B_n + H_n B_n')] + \tfrac{1}{n\sigma_0^2} c_n' H_n c_n, \quad (12)$$

where $b_{n,ii}$ are the diagonal elements of B_n, $s_{n,i} = \mathrm{E}(\epsilon_{n,i}^3)$, and $\kappa_{n,i}$ is the excess kurtosis of $\epsilon_{n,i}$ which together with H_n are defined in Section 2.3. Now by the CLT for linear-quadratic forms of Kelejian and Prucha (2001), we have,

$$\frac{\sqrt{n}\tilde{\psi}_n^*}{\tau_n(\tilde{\psi}_n^*)} \xrightarrow{D} N(0,1). \quad (13)$$

This result quickly leads to the following theorem regarding the asymptotic properties of the modified QMLE $\tilde{\lambda}_n$ of the spatial parameter λ.

Theorem 2: *Under Assumptions 1-5 and 6*, the modified QML estimator $\tilde{\lambda}_n$ is consistent and asymptotically normal, i.e., as $n \to \infty$, $\tilde{\lambda}_n \xrightarrow{p} \lambda_0$, and*

$$\sqrt{n}(\tilde{\lambda}_n - \lambda_0) \xrightarrow{D} N\big(0,\ \lim_{n\to\infty} \tau_n^2(\tilde{\lambda}_n)\big),$$

where $\tau_n^2(\tilde{\lambda}_n) = \Phi_n^{-2}\tau_n^2(\tilde{\psi}_n^*)$, $\Phi_n = \frac{1}{n}\mathrm{tr}[H_n(G_n^\circ G_n + G_n^{\circ\prime} G_n - \dot{G}_n^\circ)] + \frac{1}{n\sigma_0^2} c_n' \eta_n$, and $\dot{G}_n^\circ = \frac{d}{d\lambda} G_n^\circ = G_n^2 - \mathrm{diag}(M_n)^{-1}\mathrm{diag}(M_n G_n^2)$.

Now consider the modified QMLEs $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$ of β_0 and σ_0^2 defined below (10). Using the relation $A_n(\tilde{\lambda}_n) = A_n - (\tilde{\lambda}_n - \lambda_0) W_n$, we can write,

$$\tilde{\beta}_n = \hat{\beta}_n(\lambda_0) - (\tilde{\lambda}_n - \lambda_0)(X_n' X_n)^{-1} X_n' G_n A_n Y_n, \text{ and} \quad (14)$$

$$\tilde{\sigma}_n^2 = \hat{\sigma}_n^2(\lambda_0) - 2(\tilde{\lambda}_n - \lambda_0)\tfrac{1}{n} Y_n' W_n' M_n A_n Y_n + (\tilde{\lambda}_n - \lambda_0)^2 \tfrac{1}{n} Y_n' W_n' M_n W_n Y_n. \quad (15)$$

The asymptotic properties of $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$ are summarized in the following theorem. Recall $\eta_n = G_n X_n \beta_0$ defined in Section 2.3.

Theorem 3: *Under Assumptions 1-5 and 6*, the modified QMLEs $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$ are consistent, i.e., as $n \to \infty$, $\tilde{\beta}_n \xrightarrow{p} \beta_0$ and $\tilde{\sigma}_n^2 \xrightarrow{p} \sigma_0^2$, and further $\tilde{\beta}_n$ is asymptotically normal, i.e.,*

$$\sqrt{n}(\tilde{\beta}_n - \beta_0) \xrightarrow{D} N\big[0, \lim_{n\to\infty} (X_n' X_n)^{-1} X_n' \mathbb{A}_n X_n (X_n' X_n)^{-1}\big],$$

where $\mathbb{A}_n = n\sigma_0^2 H_n + \tau_n^2(\tilde{\lambda}_n)\eta_n\eta_n' - 2\Phi_n^{-1}(\sigma_0^{-2}\mathrm{diag}(B_n) s_n + H_n c_n)\eta_n'$, and $s_n = \mathrm{E}(\epsilon_n^3)$.[7]

Clearly, the applicability of the results of Theorems 2 and 3 for making inferences for λ or β depends on the availability of a consistent estimator of $\tau_n^2(\tilde{\psi}_n^*)$. The plug-in method based on (12) does not work due to the involvement of higher-order moments $s_{n,i}$ and $\kappa_{n,i}$.

[7]Similarly, $\sqrt{n}(\tilde{\sigma}_n^2 - \sigma_0^2) \xrightarrow{D} N\big(0, \lim_{n\to\infty} \tau_n^2(\tilde{\sigma}_n^2)\big)$, where the first-order variance of $\sqrt{n}\tilde{\sigma}_n^2$, $\tau_n^2(\tilde{\sigma}_n^2) = \frac{1}{n}\sum_{i=1}^n \mathrm{Var}(\epsilon_{n,i}^2) + \frac{4}{n^2}\sigma_0^4\tau_n^2(\tilde{\lambda}_n)\mathrm{tr}^2(H_n G_n) + \frac{4}{n^2}\mathrm{Cov}(\epsilon_n'\epsilon, \epsilon_n' B_n \epsilon_n + c_n'\epsilon_n)\mathrm{tr}(H_n G_n)\Phi_n^{-1} = O(1)$, suggesting that $\tilde{\sigma}_n^2$ is root-n consistent. However, similar to the regular QMLE, this result cannot be used for inference for σ_0^2 as the key element in the variance formula $\frac{1}{n}\sum_{i=1}^n \mathrm{Var}(\epsilon_{n,i}^2) = \frac{\sigma_0^4}{n}\sum_{i=1}^n (\kappa_{n,i} + 2h_{n,i}^2)$ cannot be consistently estimated.

3.3 Robust standard errors of the modified QML estimators

Following the discussions in Section 2.3 and Footnote 7, we focus on λ and β for robust inferences. In order to carry out inference for model parameters using the modified QML procedure, we need a consistent estimate of $\tau_n^2(\tilde{\lambda}_n)$. Given this, consistent estimates of $\tau_n^2(\tilde{\beta}_n) = (X_n'X_n)^{-1}X_n'\mathbb{A}_nX_n(X_n'X_n)^{-1}$ immediately follow. The first-order variance of the modified score as given in (12) contains second, third and fourth moments of ϵ_i which vary across i, and hence a simple White-type estimator (White, 1980) may not be suitable, which in turn makes $\tau_n^2(\tilde{\lambda}_n)$ infeasible. To overcome this difficulty, we follow the idea of Baltagi and Yang (2013b) to decompose the numerator of the modified score into a sum of uncorrelated terms, and then use the outer product of gradients (OPG) method to estimate the variance of this score function which in turn leads to a consistent estimate of $\tau^2(\tilde{\lambda}_n)$. Denote the numerator of (11) by,

$$Q_n(\epsilon_n) = \epsilon_n'B_n\epsilon_n + c_n'\epsilon_n. \tag{16}$$

Clearly, Q_n is not a sum of uncorrelated components, but can be made to be so by the technique of Baltagi and Yang (2013b). Decompose the non-stochastic matrix B_n as,

$$B_n = B_n^u + B_n^l + B_n^d, \tag{17}$$

where B_n^u, B_n^l and B_n^d are, respectively, the upper triangular, the lower triangular and the diagonal matrices of B_n. Let $\zeta_n = (B_n^{u\prime} + B_n^l)\epsilon_n$. Then, $Q_n(\epsilon_n)$ can be written as,

$$Q_n(\epsilon_n) = \textstyle\sum_{i=1}^n \epsilon_{n,i}(\zeta_{n,i} + b_{n,ii}\epsilon_{n,i} + c_{n,i}), \tag{18}$$

where $\epsilon_{n,i}$, $\zeta_{n,i}$ and $c_{n,i}$ are, respectively, the elements of ϵ_n, ζ_n and c_n. Equation (18) expresses $Q_n(\epsilon_n)$ as a sum of n uncorrelated terms $\{\epsilon_{n,i}(\zeta_{n,i} + b_{n,ii}\epsilon_{n,i} + c_{n,i})\}$, and hence its OPG gives a consistent estimate of the variance of $Q_n(\epsilon_n)$, which in turn leads to a consistent estimate of $\tau_n^2(\tilde{\psi}_n^*)$, the first-order variance of $\sqrt{n}\psi_n^*$, as:

$$\tilde{\tau}_n^2(\tilde{\psi}_n^*) = \tfrac{1}{n\tilde{\sigma}_n^4}\textstyle\sum_{i=1}^n \big(\tilde{\epsilon}_{n,i}(\tilde{\zeta}_{n,i} + \tilde{b}_{n,ii}\tilde{\epsilon}_{n,i} + \tilde{c}_{n,i})\big)^2, \tag{19}$$

where $\tilde{\epsilon}_{n,i}$ are the residuals computed from the modified QMLEs $\tilde{\theta}_n = (\tilde{\beta}_n', \tilde{\sigma}_n^2, \tilde{\lambda}_n)'$.

Let $\tilde{H}_n = \frac{1}{\tilde{\sigma}_n^2}\text{diag}(\tilde{\epsilon}_{1n}^2, \ldots, \tilde{\epsilon}_{nn}^2)$. Let $\tilde{\Phi}_n$ be Φ_n evaluated at $\tilde{\theta}_n$ and $\tilde{H}_n$, $\tilde{\eta}_n = \tilde{G}_nX_n\tilde{\beta}_n$, and $\tilde{G}_n = G_n(\tilde{\lambda}_n)$. Define the estimators of $\tau_n^2(\tilde{\lambda}_n)$ and $\tau_n^2(\tilde{\beta}_n)$ as,

$$\tilde{\tau}_n^2(\tilde{\lambda}_n) = \tilde{\Phi}_n^{-2}\tilde{\tau}_n^2(\tilde{\psi}_n^*), \text{ and} \tag{20}$$

$$\tilde{\tau}_n^2(\tilde{\beta}_n) = (X_n'X_n)^{-1}X_n'\tilde{\mathbb{A}}_nX_n(X_n'X_n)^{-1}, \tag{21}$$

where $\tilde{\mathbb{A}}_n = n\tilde{\sigma}_n^2\tilde{H}_n + \tilde{\tau}_n^2(\tilde{\lambda}_n)\tilde{\eta}_n\tilde{\eta}_n' - 2\tilde{\Phi}_n^{-1}(\tilde{\sigma}_n^{-2}\tilde{B}_n^d\tilde{s}_n + \tilde{H}_n\tilde{c}_n)\tilde{\eta}_n'$ and $\tilde{s}_n = \tilde{\epsilon}_n^3$. Note that Φ_n can simply be estimated by $-\frac{d}{d\lambda_0}\tilde{\psi}_n^*|_{\lambda_0=\tilde{\lambda}_n}$ as Φ_n is the 1st-order term of $-\text{E}(\frac{d}{d\lambda_0}\tilde{\psi}_n^*)$.

Theorem 4: *If Assumptions 1-5 and 6* hold, then we have as $n \to \infty$, $\tilde{\tau}_n^2(\tilde{\lambda}_n) - \tau_n^2(\tilde{\lambda}_n) \xrightarrow{p} 0$; and $\tilde{\tau}_n^2(\tilde{\beta}_n) - \tau_n^2(\tilde{\beta}_n) \xrightarrow{p} 0$.*

Finally, when the conditions of Theorem 1 are satisfied so the regular QMLEs are also consistent, the robust variances of $\hat{\lambda}_n$ and $\hat{\beta}_n$ can easily be obtained from the results of Theorems 2-4. Some details are as follows. Starting with the concentrated score $\tilde{\psi}_n$ given in (7), one obtains $\tau^2(\hat{\lambda}_n)$ by simply replacing G_n° by $G_n - \frac{1}{n}\text{tr}(G_n)I_n$ in (11) and (12), and in Φ_n defined in Theorem 2. Similarly, replacing G_n° by $G_n - \frac{1}{n}\text{tr}(G_n)I_n$ in $\tau_n^2(\tilde{\beta}_n)$ given in Theorem 3 leads to $\tau_n^2(\hat{\beta}_n)$. The estimates of $\tau^2(\hat{\lambda}_n)$ and $\tau_n^2(\hat{\beta}_n)$ are obtained in the same way as those of $\tau^2(\tilde{\lambda}_n)$ and $\tau_n^2(\tilde{\beta}_n)$, and their consistency can be proved similarly to the results of Theorem 4.

4. Monte Carlo Study

Extensive Monte Carlo experiments were conducted to (i) investigate the finite sample behaviour of the original QMLE $\hat{\lambda}_n$ and the modified QMLE (MQMLE) $\tilde{\lambda}_n$ proposed in this paper, and their impacts on the estimators of β and σ^2, with respect to the changes in the sample size, spatial layouts, error distributions and the model parameters when the models are heteroskedastic; and (ii) compare the QMLE and the MQMLE with the non-robust generalized method of moments estimator (GMME) of Lee (2001), the robust GMME (RGMME) and the optimal RGMME (ORGMME) of Lin and Lee (2010), two stage least squares estimator (2SLSE) of Kelejian and Prucha (1998), and the root estimator (RE) of Jin and Lee (2012). We consider cases where the original QMLE are robust against heteroskedasticity and the cases it is not.

The simulations are carried out based on the following data generation process (DGP):

$$Y_n = \lambda W_n Y_n + \iota_n \beta_0 + X_{1n}\beta_1 + X_{2n}\beta_2 + \epsilon_n,$$

where ι_n is an $n \times 1$ vector of ones corresponding to the intercept term, X_{1n} and X_{2n} are the $n \times 1$ vectors containing the values of two fixed regressors, and $\epsilon_n = \sigma H_n e_n$. The regression coefficients β is set to either $(3, 1, 1)'$ or $(.3, .1, .1)'$, σ is set to 1, λ takes values form $\{-0.5, -0.25, 0, 0.25, 0.5\}$ and n take values from $\{100, 250, 500, 1000\}$. The ways of generating the values for (X_{1n}, X_{2n}), the spatial weights matrix W_n, the heteroskedasticity measure H_n, and the idiosyncratic errors e_n are described below. Each set of Monte Carlo results is based on $1,000$ Monte Carlo samples.

Spatial Weight Matrix: We use three different spatial layouts: (i) `Circular Neighbours`, (ii) `Group Interaction` and (iii) `Queen Contiguity`. In (i), neighbours occur in the positions immediately ahead and behind a particular spatial unit. For example, for the ith spatial unit with 6 neighbours, the ith row of W_n matrix has non-zero elements in the positions: $i-3, i-2, i-1, i+1, i+2$, and $i+3$. The weights matrix we consider has 2, 4, 6, 8 and 10 neighbours with equal proportion. In (ii), neighbours occur in groups where each group member is spatially related to one another resulting in a symmetric W_n matrix. In (iii), neighbours could occur in the eight cardinal and ordinal positions of each unit. To ensure the heteroskedasticity effect does not fade as n increases (so that the regular QMLE is inconsistent), the degree of spatial dependence is fixed with respect to n. This is attained by fixing the possible group sizes in the Group Interaction scheme, and fixing the number of neighbours behind and ahead in the Circular Neighbours scheme. The degree of spatial dependence is naturally bounded in the

Queen Contiguity weight matrix. To analyse the performance of the original QMLE when it is robust against heteroskedasticity, we use Queen Contiguity scheme and the `balanced Circular Neighbours` scheme where all spatial units have 6 peers each.

Heteroskedasticity: For the `unbalanced Circular Neighbour` scheme, $h_{n,i}$ is generated as the ratio of the total number of neighbours to the average number of neighbours for each i while for the Group Interaction scheme $h_{n,i}$ is generated as the ratio of the group size to mean group size. For the balanced Circular Neighbour and the Queen Contiguity schemes, we use $h_{n,i} = n[\sum_{i=1}^{n}(|X_{1n,i}| + |X_{2n,i}|)]^{-1}(|X_{1n,i}| + |X_{2n,i}|)$.

Regressors: The regressors are generated according to `REG1:` $\{x_{1i}, x_{2i}\} \overset{iid}{\sim} N(0,1)/\sqrt{2}$. For the Group Interaction scheme, the regressors can also be generated according to `REG2:` $\{x_{1,ir}, x_{2,ir}\} \overset{iid}{\sim} (2z_r + z_{ir})/\sqrt{10}$, where $(z_r, z_{ir}) \overset{iid}{\sim} N(0,1)$, for the ith observation in the rth group, to give a case of non-iid regressors taking into account the impact of group sizes on the regressors. Both schemes give a signal-to-noise ratio of 1 when $\beta_1 = \beta_2 = \sigma = 1$.

Error Distribution: To generate the e_n component of the disturbance term, three DGPs are considered: `DGP1:` $\{e_{n,i}\}$ are iid standard normal, `DGP2:` $\{e_{n,i}\}$ are iid standardized normal mixture with 10% of values from $N(0,4)$ and the remaining from $N(0,1)$ and `DGP3:` $\{e_{n,i}\}$ iid standardized log-normal with parameters 0 and 1. Thus, the error distribution from `DGP2` is leptokurtic, and that of `DGP3` is both skewed and leptokurtic.

The GMM-type estimators are implemented by closely following Lin & Lee (2010). A GMM estimator is in general defined as a solution to the minimisation problem: $\min_{\theta\in\Theta} g_n'(\theta)a_n'a_n g_n(\theta)$ where $g_n(\theta) = \big(Q_n, P_{1n}\epsilon_n(\theta), \ldots, P_{mn}\epsilon_n(\theta)\big)'\epsilon_n(\theta)$ represents the linear and quadratic moment conditions, $Q_n = (X_n, W_n X_n)$ is the matrix of instrumental variables (IVs), and $a_n'a_n$ is the weighting matrix related to the distance function of the minimisation problem. The GMME (Kelejian & Prucha, 1999; Lee, 2001) under homoskedastic disturbances can be defined using the usual moment condition, $P_n = \big(G_n - \frac{\mathrm{tr}(G_n)}{n} I_n\big)$ and the IVs, $(G_n X_n \beta, X_n)$. For the RGMME, the P_n matrix in the moment conditions changes to $G_n - \mathrm{diag}(G_n)$. A first step GMME with $P_n = W_n$ is used to evaluate G_n. The weighting matrices of the distance functions are computed using the variance formula of the iid case using residual estimates given by the first step GMM estimate. The ORGMME is a variant of the RGMME in which the weighting matrix is robust to unknown heteroskedasticity. The ORGMME results given in the tables are computed using the RGMME as the initial estimate to compute the standard error estimates and the instruments. Finally, the 2SLSE uses the same IV matrix Q_n. Lin and Lee (2010) gives a detailed comparison of the finite sample performance of MLE, GMME, RGMME, ORGMME and 2SLSE for models with both homoskedastic and heteroskedastic errors. Our Monte Carlo experiments expand theirs by giving a detailed investigation on the effects of nonnormality, spatial layouts as well as negative values for the spatial parameter. The RE of Jin and Lee (2012) is also included.

To conserve space, only the partial results of QMLE, MQMLE, RGMME and ORGMME are reported. The full set of results are available from the authors upon request. The GMME and 2SLSE can perform very poorly. The root estimator performs equally well as the MQMLE when $|\lambda|$ is not large and n is not small but tends to give non-real roots otherwise. Tables 1-3 summarise the estimation results for λ and Tables 4-6 for β, where in each table, the Monte Carlo means, root mean square errors (rmse) and the standard errors (se) of the estimators are reported. To analyse the finite sample performance of the proposed OPG based robust standard error estimators, we also report the averaged se of the regular QMLE when it is heteroskedasticity robust and the averaged se of the MQMLE based on Theorem 4. The experiments with $\beta = (0.3, 0.1, 0.1)$ represent cases where the stochastic component is relatively more dominant than the deterministic component of the model. This allows a comparison between the QML-type estimators and the GMM-type estimators when the model suffers from relatively more severe heteroskedasticity and the IVs are weaker. The main observations made from the Monte Carlo results are summarized as follows:

(i) MQMLE of λ performs well in all cases considered, and it generally outperforms all other estimators in terms of bias and rmse.[8] Further, in cases where QMLE is consistent, MQMLE can be significantly less biased than QMLE, and is as efficient as QMLE.

(ii) RGMME and ORGMME of λ perform reasonably well when $\beta = (3, 1, 1)'$, but deteriorates significantly when $\beta = (.3, .1, .1)'$ and in this case GMME and 2SLSE can be very erratic. In contrast, MQMLE is much less affected by the magnitude of β, and is less biased and more efficient than RGMME and ORGMME more significantly when $\beta = (.3, .1, .1)'$.

(iii) RE of λ performs equally well as MQMLE when $|\lambda|$ is not big and n is not small, but otherwise tends to give imaginary roots. Thus, when one encounters a super large dataset and the QMLE or MQMLE run into computational difficulty, one may turn to RE and use its closed-form expression.

(iv) The GMM-type estimators can perform quite differently when the errors are normal as opposed to non-normal errors, especially when $\beta = (.3, .1, .1)'$. It is interesting to note that RGMME often outperforms the ORGMME.

(v) The OPG-based estimate of the robust standard errors of MQMLE of λ performs well in general with their values very close to their Monte Carlo counter parts.

(vi) Finally, the relative performance of various estimators of β is much less contrasting than that of various estimators of λ, although it can be seen that MQMLE of β is slightly more efficient than the competing RGMME and ORGMME.

5. Extension to More General Models

As discussed in the introduction and Remark 2 of Section 3.1, the modified QML estimation method can be easily extended to suit for more general models (spatial or non-spatial) where

[8] A referee points out that under homoskedasticity, the GMM estimator can be as efficient as the MLE when errors are normal, and can be more efficient than the QMLE when the errors are nonnormal. See also Lee and Liu (2010). However, under heteroskedasticity, the latter is not observed from our extensive Monte Carlo Experiments. It would be interesting to carry out a theoretical comparison on the efficiency of the heteroskedasticity robust GMM-type and QML-type estimators, but such a study is clearly beyond the scope of this paper.

there are two or more concentrated score elements that need to be modified to account for the unknown heteroskedasticity. One popular example is the so-called SARAR(1,1) model, which extends the SAR model considered above by allowing the disturbances ϵ_n to follow a heteroskedastic SAR process. In this section, we first present a full set of 'feasible' results for the SARAR(1,1) model to facilitate immediate practical applications, and then discuss further possible extensions of the proposed methods. The SARAR(1,1) model takes the form,

$$Y_n = \lambda W_{1n} Y_n + X_n \beta + \epsilon_n, \quad \epsilon_n = \rho W_{2n} \epsilon_n + v_n, \tag{22}$$

where $v_{n,i} \sim inid(0, \sigma^2 h_{n,i})$ such that and $\frac{1}{n}\sum_{i=1}^n h_{n,i} = 1$. Let $A_n(\lambda) = I_n - \lambda W_{1n}$ and $B_n(\rho) = I_n - \lambda W_{2n}$, then the concentrated Gaussian loglikelihood function for $\delta = (\lambda, \rho)'$ is,

$$\ell_n^c(\delta) = -\frac{n}{2}[\ln(2\pi) + 1] - \frac{n}{2}\ln(\hat{\sigma}_n^2(\delta)) + \ln|A_n(\lambda)| + \ln|B_n(\rho)|, \tag{23}$$

where $\hat{\sigma}_n^2(\delta) = \frac{1}{n} Y_n'(\delta) M_n(\rho) Y_n(\delta)$, $M_n(\rho) = I_n - B_n(\rho) X_n [X_n' B_n'(\rho) B_n(\rho) X_n]^{-1} X_n' B_n'(\rho)$ and $Y_n(\delta) = B_n(\rho) A_n(\lambda) Y_n$. Maximizing (23) gives the QMLE $\hat{\delta}_n$ of δ, and thus the QMLE of β as $\hat{\beta}_n \equiv \hat{\beta}_n(\hat{\delta}_n)$ where $\hat{\beta}_n(\delta) = [X_n' B_n'(\rho) B_n(\rho) X_n]^{-1} X_n' B_n'(\rho) Y_n(\delta)$, and the QMLE of σ^2 as $\hat{\sigma}_n^2 \equiv \hat{\sigma}_n^2(\hat{\delta}_n)$. The concentrated score function upon dividing by n is,

$$\tilde{\psi}_n(\delta) = \begin{cases} -\frac{1}{n}\text{tr}(G_{1n}(\lambda)) + \dfrac{Y_n'(\delta) M_n(\rho) \bar{G}_{1n}(\delta) Y_n(\delta)}{Y_n'(\delta) M_n(\rho) Y_n(\delta)}, \\ -\frac{1}{n}\text{tr}(G_{2n}(\rho)) + \dfrac{Y_n'(\delta) M_n(\rho) \bar{G}_{2n}(\rho) Y_n(\delta)}{Y_n'(\lambda) M_n(\rho) Y_n(\delta)}, \end{cases} \tag{24}$$

where $\bar{G}_{1n}(\delta) = B_n(\rho) G_{1n}(\lambda) B_n^{-1}(\rho)$, $\bar{G}_{2n}(\rho) = G_{2n}(\rho) M_n(\rho)$, $G_{1n}(\lambda) = W_{1n} A_n^{-1}(\lambda)$, and $G_{2n}(\rho) = W_{2n} B_n^{-1}(\rho)$. Using similar arguments as given in Section 3, we have, after some algebraic manipulations, the following modified concentrated score function,

$$\tilde{\psi}_n^*(\delta) = \begin{cases} \dfrac{Y_n'(\delta) M_n(\rho) \bar{G}_{1n}^\circ(\delta) Y_n(\delta)}{Y_n'(\delta) M_n(\rho) Y_n(\delta)}, \\ \dfrac{Y_n'(\delta) M_n(\rho) \bar{G}_{2n}^\circ(\rho) Y_n(\delta)}{Y_n'(\delta) M_n(\rho) Y_n(\delta)}, \end{cases} \tag{25}$$

where $\bar{G}_{rn}^\circ(\delta) = \bar{G}_{rn}(\delta) - \text{diag}(M_n(\rho))^{-1}\text{diag}[M_n(\rho)\bar{G}_{rn}(\delta)], r = 1, 2$.

The modified QMLE of δ is defined as $\tilde{\delta}_n = \arg\{\tilde{\psi}_n^*(\delta) = 0\}$, and the modified QMLEs of β and σ^2 are $\tilde{\beta}_n \equiv \hat{\beta}_n(\tilde{\delta}_n)$ and $\tilde{\sigma}_n^2 \equiv \hat{\sigma}_n^2(\tilde{\delta}_n)$. To the best of our knowledge, the three-step estimator of Kelejian and Prucha (2010) may be the only heteroskedasticity robust estimator for the SARAR(1,1) model available in the literature.[9] Thus, it would be of a great interest to investigate and compare the finite sample properties of the three-step estimator and the proposed modified QMLE estimator for the SARAR(1,1) model. For brevity, Table 7 presents a small set of Monte Carlo results that serve such purposes, and more results are available from the authors. Both the reported and unreported Monte Carlo results show that the proposed

[9] Arraiz, et al. (2010) provide some additional details for this estimator including some Monte Carlo results.

modified QMLE has an excellent finite sample performance, and it outperforms the three-step estimator of Kelejian and Prucha (2010) from a combined consideration in terms of bias, consistency and efficiency.[10]

For heteroskedasticity robust inferences based on the SARAR(1,1) model, one needs the feasible heteroskedasticity robust estimators of the asymptotic variances of $\tilde{\delta}$ and $\tilde{\beta}_n$. Under an extended set of regularity conditions and using the multivariate CLT for linear-quadratic forms of Kelejian and Prucha (2010, Appendix A), we can show that as $n \to \infty$,

$$\sqrt{n}(\tilde{\delta}_n - \delta_0) \xrightarrow{D} N\big(0, \lim_{n\to\infty} \tau_n^2(\tilde{\delta}_n)\big), \quad \text{and} \quad \tau_n^2(\tilde{\delta}_n) = \Phi_n^{-1}\tau_n^2(\tilde{\psi}_n^*)\Phi_n^{-1}, \tag{26}$$

where Φ_n equals to $-\mathrm{E}[\frac{\partial}{\partial \delta_0'}\tilde{\psi}^*(\delta_0)]$ or its first-order term, and $\tau_n^2(\tilde{\psi}_n^*)$ is the first-order terms of $\mathrm{Var}[\sqrt{n}\tilde{\psi}^*(\delta_0)]$. Both Φ_n and $\tau_n^2(\tilde{\psi}_n^*)$ possess analytical expressions but are not needed for practical applications as the former can be estimated consistently by $\tilde{\Phi}_n = -\frac{\partial}{\partial \delta_0'}\tilde{\psi}^*(\delta_0)|_{\delta_0=\tilde{\delta}_n}$, and the latter by the following OPG estimator:

$$\tilde{\tau}_n^2(\tilde{\psi}_n^*) = \textstyle\sum_{i=1}^n \tilde{\epsilon}_{n,i}^2 \tilde{\Upsilon}_{n,i}\tilde{\Upsilon}_{n,i}', \tag{27}$$

where $\tilde{\Upsilon}_{n,i} = (\tilde{\zeta}_{1n,i} + \tilde{p}_{1n,ii}\tilde{\epsilon}_{n,i} + \tilde{c}_{1n,i}, \tilde{\zeta}_{2n,i} + \tilde{p}_{2n,ii}\tilde{\epsilon}_{n,i} + \tilde{c}_{2n,i})'$, $\tilde{\zeta}_{rn} = (P_{rn}^{u\prime} + P_{rn}^{l})\tilde{\epsilon}_n$, $r = 1,2$, $\tilde{\epsilon}_n = Y(\tilde{\delta}_n) - B_n(\tilde{\rho})X_n\tilde{\beta}_n$, and P_{rn} and c_{rn} are defined in the following asymptotic representation:

$$\sqrt{n}\tilde{\psi}_n^* = \begin{cases} \frac{1}{\sqrt{n}\sigma_0^2}\big(\epsilon_n' P_{1n}\epsilon_n + c_{1n}'\epsilon_n\big) + o_p(1), \\ \frac{1}{\sqrt{n}\sigma_0^2}\big(\epsilon_n' P_{2n}\epsilon_n + c_{2n}'\epsilon_n\big) + o_p(1), \end{cases} \tag{28}$$

where $P_{rn} = M_n\bar{G}_{rn}^{\circ}$ and $c_{rn} = M_n\bar{G}_{rn}^{\circ}B_nX_n\beta_0$, $r = 1,2$, with $p_{rn,ii}$, P_{rn}^u and P_{rn}^l denoting, respectively, the diagonal elements, the upper and lower triangular matrices of P_{rn}.

With the asymptotic results for $\tilde{\delta}_n$, one can easily derive the asymptotic results for $\tilde{\beta}_n$. Under a similar set of regularity conditions, we can show that as $n \to \infty$,

$$\sqrt{n}(\tilde{\beta}_n - \beta_0) \xrightarrow{D} N\big(0, \lim_{n\to\infty}(X_n'B_n'B_nX_n)^{-1}X_n'B_n'\mathbb{A}_nB_nX_n(X_n'B_n'B_nX_n)^{-1}\big), \tag{29}$$

where $\mathbb{A}_n = n\sigma_0^2 H_n + \tau_{n,11}^2(\tilde{\delta}_n)\eta_n\eta_n' + 2\sqrt{n}(\sigma_0^{-2}P_{1n}^d s_n + H_n c_{1n}, \sigma_0^{-2}P_{2n}^d s_n + H_n c_{2n})\Phi_n^{-1}(\eta_n, 0_n)'$, $s_n = \mathrm{E}(\epsilon_n^3)$, $P_{rn}^d = \mathrm{diag}(P_{rn})$, $\eta_n = B_nG_{1n}X_n\beta_0$, $\tau_{n,11}^2(\tilde{\delta}_n)$ is the top-right corner element of $\tau_n^2(\tilde{\delta}_n)$, and 0_n is an $n \times 1$ vector of 0's. With the estimates $\tilde{\Phi}_n$ and $\tilde{\tau}_n^2(\tilde{\psi}_n^*)$ defined above, the estimates $\tilde{s}_n = \tilde{\epsilon}_n^3$ and $\tilde{H}_n = \tilde{\sigma}_n^{-2}\mathrm{diag}(\tilde{\epsilon}_n^2)$ of s_n and H_n, and the plug-in estimates for the remaining quantities, a consistent estimate for $\tau_n^2(\tilde{\beta}_n^*)$ follows.

The proposed methods can be further extended. For example, the SARAR(p,q), which contains spatial lags of order p and spatial autoregressive errors of order q, can be dealt with in a similar manner as for the SARAR(1,1) model. To have an idea on how our methods can be extended

[10] A more rigorous comparison may be interesting but beyond the scope of this paper. The robust GMM approach of Lin and Lee (2010) may lead to a more efficient estimator than does the three-step approach of Kelejian and Prucha (2010), but from Lin and Lee (2010) it is not clear how to extend their robust GMM estimation approach for the SAR to the general SARAR(1,1) model.

to the SARAR(p, q) model, note that the Gaussian likelihood takes an identical form as (24) for SARAR(1,1), except that now $A_n(\lambda) = I_n - \sum_{j=1}^{p} \lambda_j W_{1,jn}$ and $B_n(\rho) = I_n - \sum_{j=1}^{q} \rho_j W_{2,jn}$, $\lambda = \{\lambda_1, \ldots, \lambda_p\}$ and $\rho = \{\rho_1, \ldots, \rho_q\}$, see Lee and Liu (2010). Thus, the concentrated scores and their modifications can be found in a similar manner, resulting modified QMLEs for the SARAR(p, q) model that are robust against unknown heteroskedasticity.[11] Moving further, our methods can be applied to give heteroskedasticity robust estimator for the fixed effects spatial panel data model. As argued in the introduction, heteroskedasticity is common particularly in spatial models. This makes it more desirable to develop heteroskedasticity robust inference methods for these models. The methods proposed in this paper shed much light on these intriguing research problems. However, formal studies on these models, including detailed proofs of the results (26)-(29) and the proofs of consistency of the variance estimates therein, are beyond the scope of this paper, and will be pursued in future research.

6. Conclusion

This paper looks at heteroskedasticity robust QML-type estimation for spatial autoregressive (SAR) models. We provide clear conditions for the regular QMLE to be consistent even when the disturbances suffer from heteroskedasticity of unknown form. When these conditions are violated, the regular QMLE becomes inconsistent and in this case we suggest a modified QMLE by making a simple adjustment to the score function so that it becomes robust to unknown heteroskedasticity. This method is proven to work well in the simulation studies and was shown to be robust to many situations including, deteriorated signal strength as well as non-normal errors (besides the unknown heteroskedasticity). To provide inference methods robust to heteroskedasticity and nonnormality, OPG-based estimators of the variances of QMLE and modified QMLE are introduced. Monte Carlo results show that the proposed modified QMLE for the SAR model and the associated robust variance estimator work very well in finite samples.

The proposed methodology (modifying score for achieving heteroskedasticity robustness for parameter estimation and finding a suitable OPG for achieving heteroskedasticity robustness for variance estimation) has a great potential to be extended to more general models, not necessarily the spatial models, thus paving a simple way for developing heteroskedasticity robust inference methods for applied researchers.

[11]Lee and Liu (2010) proposed efficient GMM estimation of this model under homoskedasticity assumption. Badinger and Egger (2011) extend the estimation strategy of Kelejian and Prucha (2010) to give a heteroskedasticity robust three-step estimator of the SARAR(p, q) model, where some Monte Carlo results are presented under a SARAR(3,3) model and some special spatial weight matrices.

Appendix A: Some Useful Lemmas

The following lemmas are extended versions of the selected lemmas from Lee (2004), Kelejian and Prucha (2001) and Lin and Lee (2010), which are required in the proofs of the main results.

Lemma A.1: *Suppose the matrix of independent variables X_n has uniformly bounded elements, then the projection matrices $P_n = X_n(X_n'X_n)^{-1}X_n'$ and $M_n = I_n - P_n$ are uniformly bounded in both row and column sums.*

Lemma A.2: *Let A_n be an $n \times n$ matrix, uniformly bounded in both row and column sums. Then for M_n defined in Lemma A.1,*
(i) $\mathrm{tr}(A_n^m) = O(n)$ *for* $m \geq 1$,
(ii) $\mathrm{tr}(A_n'A_n) = O(n)$,
(iii) $\mathrm{tr}((M_nA_n)^m) = \mathrm{tr}(A_n^m) + O(1)$ *for* $m \geq 1$ *and*
(iv) $\mathrm{tr}((A_n'M_nA_n)^m) = \mathrm{tr}((A_n'A_n)^m) + O(1)$ *for* $m \geq 1$.
Let B_n be another $n \times n$ matrix, uniformly bounded in both row and column sums. Then,
(iv) A_nB_n *is uniformly bounded in both row and column sums,*
(v) $\mathrm{tr}(A_nB_n) = \mathrm{tr}(B_nA_n) = O(n)$ *uniformly.*

Lemma A.3 (Moments and Limiting Distribution of Quadratic Forms): *For a given process of innovations $\{\epsilon_{n,i}\}$, let $\epsilon_{n,i} \sim inid(0, \sigma_{n,i}^2)$, where $\sigma_{n,i}^2 = \sigma_0^2 h_{n,i}$, $h_{n,i} > 0$ for $i = 1, \ldots, n$ such that $\frac{1}{n}\sum_{i=1}^n h_{n,i} = 1$. Let $H_n = diag(h_{n,1}, \ldots, h_{n,n})$, B_n be an $n \times n$ matrix of elements $b_{n,ij}$, and c_n an $n \times 1$ vector of elements $c_{n,i}$. For $Q_n = \epsilon_n' B_n \epsilon_n + c_n'\epsilon_n$,*
(i) $\mathrm{E}(Q_n) = \sigma_0^2 \mathrm{tr}(H_nB_n)$ *and*
(ii) $\mathrm{Var}(Q_n) = \sum_{i=1}^n(\sigma_{n,i}^4 b_{n,ii}^2 \kappa_{n,i} + 2\sigma_{n,i}^3 b_{n,ii} c_{n,i}\gamma_{n,i}) + \sigma_0^4 tr[H_nB_n(H_nB_n + B_n'H_n)] + \sigma_0^2 c_n' H_n c_n$,
where $\gamma_{n,i}$ and $\kappa_{n,i}$ are, respectively, the skewness and excess kurtosis of $\epsilon_{n,i}$. Now, if B_n is uniformly bounded in either row or column sums then,
(iii) $\mathrm{E}(Q_n) = O(n)$,
(iv) $\mathrm{Var}(Q_n) = O(n)$,
(v) $Q_n = O_p(n)$,
(vi) $\frac{1}{n}Q_n - \frac{1}{n}\mathrm{E}(Q_n) = O_p\big(n^{-\frac{1}{2}}\big)$ *and*
(vii) $\mathrm{Var}(\frac{1}{n}Q_n) = O(n^{-1})$.
Further, if B_n is uniformly bounded in both row and column sums and Assumption 4 holds then,
(viii) $\frac{Q_n - \mathrm{E}(Q_n)}{\sqrt{\mathrm{Var}(Q_n)}} \xrightarrow{D} N(0, 1)$.

Appendix B: Proofs of Theorems and Corollaries

Proof of Theorem 1: We only prove the consistency of $\hat{\lambda}_n$ as the consistency of $\hat{\beta}_n$ and $\hat{\sigma}_n^2$ immediately follows from identities similar to (14) and (15). Define $\bar{\ell}_n^c(\lambda) = \max_{\beta,\sigma^2} \mathrm{E}[\ell_n(\theta)]$. By Theorem 5.7 of van der Vaart (1998), it amounts to show, (a) identification uniqueness condition: $\sup_{\lambda:d(\lambda,\lambda_0)\geq\epsilon} \frac{1}{n}[\bar{\ell}_n^c(\lambda) - \bar{\ell}_n^c(\lambda_0)] < 0$ for any $\epsilon > 0$ and a distance measure $d(\lambda, \lambda_0)$ and (b) uniform convergence: $\frac{1}{n}[\ell_n^c(\lambda) - \bar{\ell}_n^c(\lambda)] \xrightarrow{p} 0$ uniformly in $\lambda \in \Lambda$.

It is easy to see that $\bar{\ell}_n^c(\lambda) = -\frac{n}{2}(\ln(2\pi)+1) - \frac{n}{2}\ln(\bar{\sigma}_n^2(\lambda)) + \ln|A_n(\lambda)|$, where $\bar{\sigma}_n^2(\lambda) = \frac{1}{n}\big[(\lambda_0-\lambda_n)^2\eta_n' M_n\eta_n + \sigma_0^2\mathrm{tr}[H_n A_n'^{-1}A_n'(\lambda)A_n(\lambda)A_n^{-1}]\big]$. Recall $\ell_n^c(\lambda)$ defined in (3).

Condition (a)**:** Observe that $\bar{\sigma}_n^2(\lambda_0) = \sigma_0^2$, then,

$$\begin{aligned} & \lim_{n\to\infty}\tfrac{1}{n}\big[\bar{\ell}_n^c(\lambda) - \bar{\ell}_n^c(\lambda_0)\big] \\ = \; & \lim_{n\to\infty}\big[\tfrac{1}{2n}(\log|A_n'(\lambda)A_n(\lambda)| - \log|A_n'A_n|) + \tfrac{1}{2n}(\log|\sigma_n^{-2}(\lambda)I_n| - \log|\sigma_0^{-2}I_n|)\big] \\ \neq \; & 0 \text{ for } \lambda\neq\lambda_0, \text{ by Assumption 6.} \end{aligned}$$

Next, note that $p_n(\theta_0) = \exp[\ell_n(\theta_0)]$ is the *quasi* joint pdf of ϵ_n, which is $N(0,\sigma^2 I_n)$. Let $p_n^0(\theta_0)$ be the *true* joint pdf of $\epsilon_n \sim (0,\sigma^2 H_n)$. Let E^q denote the expectation with respect to $p_n(\theta_0)$, to differentiate from the usual notation E that corresponds to $p_n^0(\theta_0)$.

Now consider $\epsilon_n(\beta,\lambda) = A_n(\lambda)Y_n - X_n\beta = B_n(\lambda)\epsilon_n + b_n(\beta,\lambda)$, where $B_n(\lambda) = A_n(\lambda)A_n^{-1}$ and $b_n(\beta,\lambda) = A_n(\lambda)A_n^{-1}X_n\beta_0 - X_n\beta$. Then, with $\ell_n(\theta)$ given in (2), we have

$$\begin{aligned} \mathrm{E}^q[\ell_n(\theta_0)] &= -\tfrac{n}{2}\ln(2\pi\sigma^2) + \ln|A_n| - \tfrac{n}{2}, \\ \mathrm{E}[\ell_n(\theta_0)] &= -\tfrac{n}{2}\ln(2\pi\sigma^2) + \ln|A_n| - \tfrac{n}{2}, \text{ as } \tfrac{1}{n}\textstyle\sum_{i=1}^n h_{n,i} = 1 \\ \mathrm{E}^q[\ell_n(\theta)] &= -\tfrac{n}{2}\ln(2\pi\sigma^2) + \ln|A_n(\lambda)| - \tfrac{1}{2\sigma^2}[\sigma_0^2\mathrm{tr}(B_n'(\lambda)B_n(\lambda)) + b_n'(\beta,\lambda)b_n(\beta,\lambda)], \\ \mathrm{E}[\ell_n(\theta)] &= -\tfrac{n}{2}\ln(2\pi\sigma^2) + \ln|A_n(\lambda)| - \tfrac{1}{2\sigma^2}[\sigma_0^2\mathrm{tr}(H_nB_n'(\lambda)B_n(\lambda)) + b_n'(\beta,\lambda)b_n(\beta,\lambda)], \end{aligned}$$

where we have used the identities, $B_n(\lambda_0) = I_n$ and $b_n(\beta_0,\lambda_0) = 0$. Now using the identities $A_n(\lambda) = A_n + (\lambda_0-\lambda)W_n$ and $B_n(\lambda) = I_n + (\lambda_0-\lambda)G_n$, we have,

$$\begin{aligned} & \mathrm{E}[\ell_n(\theta)] - \mathrm{E}^q[\ell_n(\theta)] \\ & = 2(\lambda_0-\lambda)[\mathrm{tr}(H_nG_n) - tr(G_n)] + (\lambda_0-\lambda)^2[\mathrm{tr}(H_nG_n'G_n) - \mathrm{tr}(G_n'G_n)] = o(1), \end{aligned}$$

where the last equality holds by assumptions $\mathrm{Cov}(g_n,h_n) = o(1)$ and $\mathrm{Cov}(q_n,h_n) = o(1)$.

Now by Jensen's inequality, $0 = \log\mathrm{E}^q\big(\frac{p_n(\theta)}{p_n(\theta_0)}\big) \geq \mathrm{E}^q\big[\log\big(\frac{p_n(\theta)}{p_n(\theta_0)}\big)\big]$, and the above results, we conclude that $\mathrm{E}\big[\log\big(\frac{p_n(\theta)}{p_n(\theta_0)}\big)\big] \leq 0$ or $\mathrm{E}[\log p_n(\theta)] \leq \mathrm{E}[\log p_n(\theta_0)]$, for large enough n. Thus,

$$\bar{\ell}_n(\lambda) = \max_{\beta,\sigma^2}\mathrm{E}[\log p_n(\theta)] \leq \max_{\beta,\sigma^2}\mathrm{E}[\log p_n(\theta_0)] = \mathrm{E}[\log p_n(\theta_0)] = \bar{\ell}_n(\lambda_0), \quad \text{for } \lambda\neq\lambda_0,$$

and n large enough. The identification uniqueness condition thus follows.

Condition (b)**:** Note that $\frac{1}{n}[\ell_n^c(\lambda) - \bar{\ell}_n^c(\lambda)] = -\frac{1}{2}[\log(\hat{\sigma}_n^2(\lambda)) - \log(\bar{\sigma}_n^2(\lambda))]$. By the mean value theorem, $\log(\hat{\sigma}_n^2(\lambda)) - \log(\bar{\sigma}_n^2(\lambda)) = \frac{1}{\dot{\sigma}_n^2(\lambda)}[\hat{\sigma}_n^2(\lambda) - \bar{\sigma}_n^2(\lambda)]$, where $\dot{\sigma}_n^2(\lambda)$ lies between $\hat{\sigma}_n^2(\lambda)$ and $\bar{\sigma}_n^2(\lambda)$. Using $M_nA_n(\lambda)Y_n = (\lambda_0-\lambda)M_n\eta_n + M_nA_n(\lambda)A_n^{-1}\epsilon_n$ we can write,

$$\hat{\sigma}_n^2(\lambda) = (\lambda_0-\lambda)^2\tfrac{1}{n}\eta_n'M_n\eta_n + 2(\lambda_0-\lambda)T_{1n}(\lambda) + T_{2n}(\lambda), \tag{B-1}$$

where $T_{1n}(\lambda) = \frac{1}{n}\eta_n'M_nA_n(\lambda)A_n^{-1}\epsilon$ and $T_{2n}(\lambda) = \frac{1}{n}\epsilon_n'A_n^{-1}A_n'(\lambda)M_nA_n(\lambda)A_n^{-1}\epsilon_n$.

Using $A_n(\lambda) = A_n + (\lambda_0-\lambda)W_n$, we have, $T_{1n}(\lambda) = o_p(1)$ uniformly. Further, $T_{2n}(\lambda) = \frac{1}{n}\epsilon_n'A_n^{-1}A_n'(\lambda)A_n(\lambda)A_n^{-1}\epsilon_n + o_p(1)$, since, $\frac{1}{n}\epsilon_n'A_n^{-1}A_n'(\lambda)P_nA_n(\lambda)A_n^{-1}\epsilon_n = \frac{1}{n}[\epsilon_n'P_n\epsilon + 2\epsilon_n'G_n'P_n\epsilon_n + \epsilon_n'G_n'P_nG_n\epsilon_n] = o_p(1)$ uniformly, using the condition $\mathrm{Cov}(h_n,g_n) = o(1)$. Now, Lemmas A.1 -

A.3 imply, $\frac{1}{n^2}\text{Var}(\epsilon_n' A_n^{-1} A_n'(\lambda) A_n(\lambda) A_n^{-1}\epsilon_n) = o(1)$. Then, together with Chebyshev inequality, $T_{2n}(\lambda) - \sigma_0^2 \frac{1}{n}\text{tr}[H_n A_n'^{-1} A_n'(\lambda) A_n(\lambda) A_n^{-1}] = o_p(1)$, uniformly for $\lambda \in \Lambda$.

It left to show $\sigma_n^2(\lambda)$ (defined in Assumption 6 and the main part of $\bar{\sigma}_n^2(\lambda)$) is uniformly bounded away from zero. Suppose $\sigma_n^2(\lambda)$ is not uniformly bounded away from zero. Then $\exists\{\lambda_n\} \subset \Lambda$ such that $\sigma_n^2(\lambda_n) \to 0$. Consider the model with $\beta_0 = 0$. The Gaussian log-likelihood is $\ell_{t,n}(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \log|A_n(\lambda)| - \frac{1}{2\sigma^2} Y_n' A_n'(\lambda) A_n(\lambda) Y_n$ and $\bar{\ell}_{t,n}(\lambda) = \max_{\sigma^2} \text{E}[\ell_{t,n}(\theta)]$. By Jensen's inequality, we have $\bar{\ell}_{t,n}(\lambda) \leq \max_{\sigma^2} \text{E}[\ell_{t,n}(\theta_0)] = \bar{\ell}_{t,n}(\lambda_0)$. Then together with Lemma A.2, we have $\frac{1}{n}[\bar{\ell}_{t,n}(\lambda) - \bar{\ell}_{t,n}(\lambda_0)] \leq 0$, and $-\frac{n}{2}\log(\sigma_n^2(\lambda)) \leq -\frac{n}{2}\log(\sigma_0^2) + \frac{1}{n}(\log|A_n(\lambda_0)| - \log|A_n(\lambda)|) = O(1)$. That is, $-\frac{n}{2}\log(\sigma_n^2(\lambda))$ is bounded from above which is a contradiction. Hence, $\sigma_n^2(\lambda)$ is bounded away from zero uniformly, and $\frac{n}{2}\log(\sigma_n^2(\lambda))$ is well defined $\forall\lambda \in \Lambda$.

Collecting all these results we have, $\sup_{\lambda\in\Lambda}\frac{1}{n}|[\ell_n^c(\lambda) - \bar{\ell}_n^c(\lambda)]| = o_p(1)$, completing the proof of consistency part.

To prove the asymptotic normality, first note that $\text{tr}(H_n) = n$. By the mean value theorem, $\sqrt{n}(\hat{\theta}_n - \theta_0) = -\big[\frac{1}{n}\frac{\partial^2}{\partial\theta\partial\theta'}\ell_n(\tilde{\theta})\big]^{-1}\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\ell_n(\theta_0)$, where $\tilde{\theta}_n$ lies elementwise between $\hat{\theta}_n$ and θ_0. By Assumptions 1-6, the condition $\text{Cov}(g_n, h_n) = o(n^{-1/2})$, and the CLT for vector linear-quadratic forms of Kelejian and Prucha (2010, p. 63), we have $\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\ell_n(\theta_0) \xrightarrow{D} N(0, \Sigma)$, where Σ is defined in the theorem.

Let $\mathcal{H}_n(\theta) = \frac{\partial^2}{\partial\theta\partial\theta'}\ell_n(\theta)$. It left to show (i) $\frac{1}{n}\mathcal{H}_n(\tilde{\theta}_n) - \mathcal{H}_n = o_p(1)$ and (ii) $\mathcal{H}_n - \mathbb{I}_n = o_p(1)$.

Condition (i)**:** By Assumptions 3-5 and the assumption that $\text{Cov}(h_n, g_n) = o(1)$ stated in the theorem, Lemma A.2-A.3, $\tilde{\theta}_n - \theta_0 = o_p(1)$, $\epsilon_n(\tilde{\beta}_n, \tilde{\lambda}_n) = X_n(\beta_0 - \tilde{\beta}_n) + (\lambda_0 - \tilde{\lambda}_n)W_n Y_n + \epsilon_n$ and $\frac{1}{n}\epsilon_n'(\tilde{\beta}_n, \tilde{\lambda}_n)\epsilon_n(\tilde{\beta}_n, \tilde{\lambda}_n) = \frac{1}{n}\epsilon_n'\epsilon_n + o_p(1)$, we have,

$$
\begin{aligned}
\mathcal{H}_{n,\beta\beta}(\tilde{\theta}_n) - \mathcal{H}_{n,\beta\beta} &= \big(\tfrac{1}{\sigma_0^2} - \tfrac{1}{\tilde{\sigma}_n^2}\big)\tfrac{1}{n}X_n'X_n = o_p(1),\\
\mathcal{H}_{n,\sigma^2\beta}(\tilde{\theta}_n) - \mathcal{H}_{n,\sigma^2\beta} &= \tfrac{1}{\sigma_0^4 n}\epsilon_n'X_n - \tfrac{1}{\tilde{\sigma}_0^4 n}(X_n(\beta_0 - \tilde{\beta}_n) + (\lambda_0 - \tilde{\lambda}_n)W_nY_n + \epsilon_n)'X_n = o_p(1),\\
\mathcal{H}_{n,\sigma^2\sigma^2}(\tilde{\theta}_n) - \mathcal{H}_{n,\sigma^2\sigma^2} &= \tfrac{1}{n}\big(\tfrac{1}{\sigma_0^6}\epsilon_n'\epsilon_n - \tfrac{1}{\tilde{\sigma}_n^6}\epsilon_n'(\tilde{\delta}_n)\epsilon_n(\tilde{\delta}_n)\big) - \tfrac{1}{2}\big(\tfrac{1}{\sigma_0^4} - \tfrac{1}{\tilde{\sigma}_n^4}\big) = o_p(1),\\
\mathcal{H}_{n,\lambda\beta}(\tilde{\theta}_n) - \mathcal{H}_{n,\lambda\beta} &= \big(\tfrac{1}{\sigma_0^2} - \tfrac{1}{\tilde{\sigma}_n^2}\big)\tfrac{1}{n}Y_n'W_n'X_n = o_p(1),\\
\mathcal{H}_{n,\lambda\sigma^2}(\tilde{\theta}_n) - \mathcal{H}_{n,\lambda\sigma^2} &= \tfrac{1}{\sigma_0^4 n}Y_n'W_n'\epsilon_n - \tfrac{1}{\tilde{\sigma}_0^4 n}Y_n'W_n'(X_n(\beta_0 - \tilde{\beta}_n) + (\lambda_0 - \tilde{\lambda}_n)W_nY_n + \epsilon_n) = o_p(1),\\
\mathcal{H}_{n,\lambda\lambda}(\tilde{\theta}_n) - \mathcal{H}_{n,\lambda\lambda} &= \big(\tfrac{1}{\sigma_0^2} - \tfrac{1}{\tilde{\sigma}_n^2}\big)\tfrac{1}{n}Y_n'W_n'W_nY_n + \tfrac{1}{n}\text{tr}(G_n^2) - \text{tr}(G_n^2(\tilde{\lambda}_n)) = o_p(1),
\end{aligned}
$$

where the last equality holds since $\text{tr}(G_n^2) - \text{tr}(G_n^2(\tilde{\lambda}_n)) = 2\text{tr}(G_n^2(\bar{\lambda}_n))(\lambda_0 - \tilde{\lambda}_n)$ by the mean value theorem for some $\bar{\lambda}_n$ between λ_0 and $\tilde{\lambda}_n$.

Condition (ii)**:** Given $\text{E}(\epsilon_n'\epsilon_n) = \sigma_0^2\text{tr}(H_n)$, $\text{E}(\epsilon_n'G_n\epsilon_n) = \sigma_0^2\text{tr}(H_nG_n)$, $\text{E}(\epsilon_n'G_n'G_n\epsilon_n) = \sigma_0^2\text{tr}(H_nG_n'G_n)$ and Lemma A.1-A.3, we have, $\text{Var}(\frac{1}{n}\epsilon_n'\epsilon_n) = \frac{1}{n^2}\big(\text{E}(\epsilon_{n,i}^4) - \sigma_0^4\text{tr}(H_n^2)\big) = o(1)$, $\text{Var}(\frac{1}{n}\epsilon_n'G_n\epsilon_n) = \frac{1}{n^2}\sum_{i=1}^n g_{n,ii}^2[\text{E}(\epsilon_{n,i}^4) - 3\sigma_0^4h_i^2] + \frac{1}{n^2}\sigma_0^4\text{tr}[H_nG_n(G_n'H_n + H_nG_n)] = o(1)$ and similarly $\text{Var}(\frac{1}{n}\epsilon_n'G_n'G_n\epsilon_n) = o_p(1)$. By these results and Chebyshev inequality, we have,

$$
\begin{aligned}
\mathcal{H}_{n,\beta\beta} - \mathbb{I}_{n,\beta\beta} &= 0,\\
\mathcal{H}_{n,\sigma^2\beta} - \mathbb{I}_{n,\sigma^2\beta} &= O_p(\tfrac{1}{\sqrt{n}}) = o_p(1),\\
\mathcal{H}_{n,\sigma^2\sigma^2} - \mathbb{I}_{n,\sigma^2\sigma^2} &= \tfrac{1}{\sigma_0^6}\big(\tfrac{\epsilon_n'\epsilon_n}{n} - \sigma_0^2\big) = o_p(1),\\
\mathcal{H}_{n,\lambda\beta} - \mathbb{I}_{n,\lambda\beta} &= \tfrac{1}{n}X_n'G_n\epsilon_n = O_p(\tfrac{1}{\sqrt{n}}) = o_p(1),
\end{aligned}
$$

$$\mathcal{H}_{n,\lambda\sigma^2} - \mathbb{I}_{n,\lambda\sigma^2} = \tfrac{1}{\sigma_0^4 n}\epsilon_n' G_n \epsilon_n - \tfrac{1}{\sigma_0^2 n}\mathrm{tr}(H_n G_n) + O_p(\tfrac{1}{\sqrt{n}}) = o_p(1) \text{ and}$$
$$\mathcal{H}_{n,\lambda\lambda} - \mathbb{I}_{n,\lambda\lambda} = \tfrac{1}{n}\epsilon_n' G_n' G_n \epsilon_n - \tfrac{1}{n}\mathrm{tr}(H_n G_n' G_n) + O_p(\tfrac{1}{\sqrt{n}}) = o_p(1).$$

Proof of Theorem 2: Let $\mathrm{E}(\tilde{\psi}_n^*(\lambda)) = \bar{\psi}^*(\lambda)$. By Theorem 5.9 of van der Vaart (1998), the proof of consistency of $\tilde{\lambda}_n$ requires (a) Convergence: $\sup_{\lambda\in\Lambda}|\tilde{\psi}_n^*(\lambda) - \bar{\psi}^*(\lambda)| = o_p(1)$ and (b) Identification uniqueness: for $\epsilon > 0$, $\inf_{\lambda:d(\lambda,\lambda_0)\geq\epsilon}|\bar{\psi}^*(\lambda)| > 0 = |\bar{\psi}^*(\lambda_0)|$.

The proof of Theorem 1 implies that $\hat{\sigma}_n^2(\lambda)$ is bounded away from 0 with probability one for large enough n. Thus, the modified QML estimator $\tilde{\lambda}_n = \arg\{\tilde{\psi}_n^*(\lambda) = 0\}$ is equivalently defined as $\tilde{\lambda}_n = \arg\{Y_n' A_n'(\lambda) M_n G_n^\circ(\lambda) A_n(\lambda) Y_n = 0\}$, suggesting that we can work purely with the numerator $T_n(\lambda) = Y_n' A_n'(\lambda) M_n G_n^\circ(\lambda) A_n(\lambda) Y_n$ of $\tilde{\psi}_n^*(\lambda)$ to establish consistency. Note $T_n(\lambda) = Y_n' A_n'(\lambda) M_n G_n(\lambda) A_n(\lambda) Y_n - Y_n' A_n'(\lambda) M_n \mathrm{diag}(M_n)^{-1}\mathrm{diag}(M_n G_n(\lambda)) A_n(\lambda) Y_n \equiv T_{1n}(\lambda) - T_{2n}(\lambda)$.

Condition (a): By $M_n X_n = 0$, $A_n(\lambda) = A_n + (\lambda_0 - \lambda) W_n$ and $G_n A_n = W_n = G_n(\lambda) A_n(\lambda)$,

$$\begin{aligned} T_{1n}(\lambda) &= Y_n' A_n'(\lambda) M_n G_n(\lambda) A_n(\lambda) Y_n \\ &= Y_n' A_n' M_n G_n A_n Y_n + (\lambda_0 - \lambda) Y_n' A_n' G_n' M_n G_n A_n Y_n \\ &= \epsilon_n' M_n G_n (X_n\beta_0 + \epsilon_n) + (\lambda_0 - \lambda)(X_n\beta_0 + \epsilon_n)' G_n' M_n G_n (X_n\beta_0 + \epsilon_n). \quad \text{(B-2)} \end{aligned}$$

Then, $\mathrm{E}(T_{1n}(\lambda)) = (\lambda_0 - \lambda)\beta_0' X_n G_n' M_n G_n X_n \beta_0 + \sigma_0^2 \mathrm{tr}(H_n M_n G_n) + \sigma_0^2(\lambda_0 - \lambda)\mathrm{tr}(H_n G_n' M_n G_n)$. By Lemma A.3 and Assumptions 5 and 6, we have $\frac{1}{n}[T_{1n}(\lambda) - \mathrm{E}(T_{1n}(\lambda))] = o_p(1)$. Now, as M_n appeared in T_{2n} is a projection matrix, by Lemma A.2, similar arguments as for $T_{1n}(\lambda)$ lead to $\frac{1}{n}[T_{2n}(\lambda) - \mathrm{E}(T_{2n}(\lambda))] = o_p(1)$. Thus, $\frac{1}{n}\{T_n(\lambda) - \mathrm{E}[T_n(\lambda)]\} = o_p(1)$.

Condition (b): First, we have $\mathrm{E}[T_n(\lambda_0)] = 0$, as $\mathrm{tr}[H_n M_n \mathrm{diag}(M)^{-1}\mathrm{diag}(M_n G_n)] = \mathrm{tr}[\mathrm{diag}(H_n M_n \mathrm{diag}(M)^{-1})\mathrm{diag}(M_n G_n)] = \mathrm{tr}(H_n M_n G_n)$. Now,

$$\mathrm{E}[T_n(\lambda)] = \beta_0' X_n' A_n'^{-1} A_n'(\lambda) M_n G_n^\circ(\lambda) A_n(\lambda) A_n^{-1} X_n \beta_0 + \sigma_0^2 \mathrm{tr}(H_n A_n'^{-1} A_n'(\lambda) M_n G_n^\circ(\lambda) A_n(\lambda) A_n^{-1}).$$

By Assumption 6* and Lemma A.2, $\mathrm{E}[T_n(\lambda)] \neq 0$, for any $\lambda \neq \lambda_0$. It follows that the conditions of Theorem 5.9 of van der Vaart (1998) hold, and thus the consistency of $\tilde{\lambda}_n$ follows.

To prove asymptotic normality, we have, by the mean value theorem,

$$0 = \sqrt{n}\tilde{\psi}_n^*(\tilde{\lambda}_n) = \sqrt{n}\tilde{\psi}_n^*(\lambda_0) + \tfrac{d}{d\lambda}\tilde{\psi}_n^*(\bar{\lambda}_n)\sqrt{n}(\tilde{\lambda}_n - \lambda_0), \quad \text{(B-3)}$$

where $\bar{\lambda}_n$ lies between $\tilde{\lambda}_n$ and λ_0. It suffices to show that (i) $\frac{d}{d\lambda}\tilde{\psi}_n^*(\bar{\lambda}_n) - \frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0) = o_p(1)$, (ii) $\frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0) - \mathrm{E}\big(\frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0)\big) = o_p(1)$, and (iii) $\mathrm{E}\big(\frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0)\big) \neq 0$ for large enough n. Note,

$$\begin{aligned} &\tfrac{d}{d\lambda}\tilde{\psi}_n^*(\lambda) = \tfrac{1}{n\hat{\sigma}_n^2(\lambda)} Y_n' A_n'(\lambda) \dot{G}_n^{\circ\prime}(\lambda) M_n A_n(\lambda) Y_n - \tfrac{1}{n\hat{\sigma}_n^2(\lambda)} Y_n' W_n' G_n^{\circ\prime}(\lambda) M_n A_n(\lambda) Y_n \\ &- \tfrac{1}{n\hat{\sigma}_n^2(\lambda)} Y_n' A_n'(\lambda) G_n^{\circ\prime}(\lambda) M_n W_n Y_n + \tfrac{2}{n^2\hat{\sigma}_n^4(\lambda)} Y_n' A_n'(\lambda) G_n^{\circ\prime}(\lambda) M_n A_n(\lambda) Y_n \cdot Y_n' W_n' M_n A_n(\lambda) Y_n, \end{aligned}$$

where $\dot{G}_n^\circ(\lambda) = \frac{d}{d\lambda} G_n^\circ(\lambda) = G_n^2(\lambda) - \mathrm{diag}(M_n)^{-1}\mathrm{diag}(M_n G_n^2(\lambda))$.

Condition (i): $\frac{1}{n} Y_n' W_n' M_n A_n(\bar{\lambda}_n) Y_n = \frac{1}{n} Y_n' W_n' M_n A_n Y_n + \frac{1}{n}(\lambda_0 - \bar{\lambda}_n) Y_n' W_n' M_n W_n Y_n = \frac{1}{n} Y_n' W_n' M_n A_n Y_n + o_p(1)$. Next, by Assumptions 4 and 5 and continuous mapping theorem,

$G_n^\circ(\bar{\lambda}_n) = G_n^\circ + o_p(1)$ and $\dot{G}_n^\circ(\bar{\lambda}_n) = \dot{G}_n^\circ + o_p(1)$. These lead to $\frac{1}{n}Y_n'A_n'(\bar{\lambda}_n)G_n^{\circ\prime}(\bar{\lambda}_n)M_nA_n(\bar{\lambda}_n)Y_n = \frac{1}{n}Y_n'A_n'G_n^{\circ\prime}M_nA_nY_n + o_p(1)$, and $\frac{1}{n}Y_n'A_n'(\bar{\lambda}_n)\dot{G}_n^{\circ\prime}(\bar{\lambda}_n)M_nA_n(\bar{\lambda}_n)Y_n = \frac{1}{n}Y_n'A_n'\dot{G}_n^{\circ\prime}M_nA_nY_n + o_p(1)$, after some algebra. Similarly, $\frac{1}{n}Y_n'W_n'G_n^{\circ\prime}(\bar{\lambda}_n)M_nA_n(\bar{\lambda}_n)Y_n = \frac{1}{n}Y_n'W_n'G_n^{\circ\prime}M_nA_nY_n + o_p(1)$, and $\frac{1}{n}Y_n'A_n'(\bar{\lambda}_n)G_n^{\circ\prime}(\bar{\lambda}_n)M_nW_nY_n = \frac{1}{n}Y_n'A_n'G_n^{\circ\prime}M_nW_nY_n + o_p(1)$. Collecting these results and observing $\hat{\sigma}_n^2(\bar{\lambda}_n) = \hat{\sigma}_n^2(\lambda_0) + o_p(1)$, we have $\frac{d}{d\lambda}\tilde{\psi}_n^*(\bar{\lambda}_n) - \frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0) = o_p(1)$.

Condition (ii)**:** Note that,

$$\begin{aligned} \tfrac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0) &= \tfrac{1}{n\sigma_0^2}Y_n'A_n'\dot{G}_n^{\circ\prime}M_nA_nY_n - \tfrac{1}{n\sigma_0^2}Y_nW_n'G_n^{\circ\prime}M_nA_nY_n - \tfrac{1}{n\sigma_0^2}Y_nA_n'G_n^{\circ\prime}M_nW_nY_n \\ &\quad + \tfrac{2}{n^2\sigma_0^4}(Y_n'A_n'G_n^{\circ\prime}M_nA_nY_n)\cdot(Y_n'W_n'M_nA_nY_n) + o_p(1) \equiv \textstyle\sum_{i=1}^4 T_{in} + o_p(1). \end{aligned}$$

Using $M_nA_nY_n = M_n\epsilon_n$ and the result $\frac{1}{n}a_n'\epsilon_n = o_p(1)$ for a vector a_n of uniformly bounded elements, we can readily verify that $T_{1n} = \frac{1}{n\sigma_0^2}\epsilon_n'\dot{G}_n^{\circ\prime}\epsilon_n + o_p(1)$, $T_{2n} = -\frac{1}{n\sigma_0^2}\epsilon_n'G_n^\circ G_n\epsilon_n + o_p(1)$, $T_{3n} = -\frac{1}{n\sigma_0^2}(c_n'\eta_n + \epsilon_n'G_n^{\circ\prime}G_n\epsilon_n) + o_p(1)$, and $T_{4n} = o_p(1)$, by Lemma A.2. It follows that

$$-\mathrm{E}\big[\tfrac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0)\big] = \tfrac{1}{n}\mathrm{tr}[H_n(G_n^\circ G_n + G_n^{\circ\prime}G_n - \dot{G}_n^\circ)] + \tfrac{1}{n\sigma_0^2}c_n'\eta_n + o(1) = \Phi_n + o(1),$$

and that $\frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0) - \mathrm{E}\big[\frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0)\big] = o_p(1)$.

Condition (iii)**:** By Assumptions 3-6 and Lemmas A.2 and A.3, it is easy to see that $\Phi_n \neq 0$ for large enough n, and thus $\mathrm{E}\big(\frac{d}{d\lambda}\tilde{\psi}_n^*(\lambda_0)\big) \neq 0$ for large enough n.

With (13), and (i)-(iii) proved above, the asymptotic normality result of Theorem 2 follows.

Proof of Theorem 3: Recall $\tilde{\beta}_n = (X_n'X_n)^{-1}X_n'A_n(\tilde{\lambda}_n)Y_n$. We have,

$$\sqrt{n}(\tilde{\beta}_n - \beta_0) = (\tfrac{1}{n}X_n'X_n)^{-1}\tfrac{1}{\sqrt{n}}X_n'\epsilon_n - \sqrt{n}(\tilde{\lambda}_n - \lambda_0)(\tfrac{1}{n}X_n'X_n)^{-1}\tfrac{1}{n}X_n'\eta_n + O_p\big(\tfrac{1}{\sqrt{n}}\big). \tag{B-4}$$

The proof of the asymptotic normality of $\tilde{\lambda}_n$ in Theorem 2 and the asymptotic representation for $\sqrt{n}\tilde{\psi}_n^*$ given in (11) imply that

$$\sqrt{n}(\tilde{\lambda}_n - \lambda_0) = \Phi_n^{-1}\sqrt{n}\tilde{\psi}_n^* + o_p(1) = (\sqrt{n}\sigma_0^2\Phi_n)^{-1}(\epsilon_n'B_n\epsilon_n + c_n'\epsilon_n) + o_p(1). \tag{B-5}$$

This shows that each component of $\sqrt{n}(\tilde{\beta}_n - \beta_0)$ is a linear-quadratic form in ϵ_n. Thus, Cramèr-Wold device and the CLT for linear-quadratic form of Kelejian and Prucha (2001) lead to the asymptotic normality of $\sqrt{n}(\tilde{\beta}_n - \beta_0)$. Clearly, the asymptotic mean of $\sqrt{n}(\tilde{\beta}_n - \beta_0)$ is zero and the first-order variance of it can be easily found using (B-4) and (B-5):

$$\begin{aligned} \tau^2(\tilde{\beta}_n) &= (X_n'X_n)^{-1}X_n'\mathrm{Var}(\epsilon_n)X_n(X_n'X_n)^{-1} + \tau^2(\tilde{\lambda}_n)(X_n'X_n)^{-1}X_n'\eta_n\eta_n'X_n(X_n'X_n)^{-1} \\ &\quad -2(\sigma_0^2\Phi_n)^{-1}(X_n'X_n)^{-1}X_n'\mathrm{Cov}(\epsilon_n, \epsilon_n'B_n\epsilon_n + c_n'\epsilon_n)\eta_n'X_n(X_n'X_n)^{-1} \\ &= (X_n'X_n)^{-1}X_n'\mathbb{A}_nX_n(X_n'X_n)^{-1}, \end{aligned}$$

where $\mathbb{A}_n = n\sigma_0^2H_n + \tau_n^2(\tilde{\lambda}_n)\eta_n\eta_n' - 2\Phi_n^{-1}(\sigma_0^{-2}\mathrm{diag}(B_n)s_n + H_nc_n)\eta_n'$, and $s_n = \mathrm{E}(\epsilon_n^3)$.

The limiting distribution of $\sqrt{n}(\tilde{\sigma}_n^2 - \sigma_0^2)$ can be found in a similar manner from

$$\begin{aligned}\sqrt{n}(\tilde{\sigma}_n^2 - \sigma_0^2) &= \sqrt{n}[\tfrac{1}{n}Y_n' A_n'(\tilde{\lambda}_n) M_n A_n(\tilde{\lambda}_n) Y_n - \sigma_0^2] \\ &= \tfrac{1}{\sqrt{n}}(\epsilon_n'\epsilon_n - n\sigma_0^2) + 2\sqrt{n}(\tilde{\lambda}_n - \lambda_0)\tfrac{1}{n}\sigma_0^2 \text{tr}(H_n G_n) + o_p(1),\end{aligned}$$

which has a limiting mean of zero and first-order variance:

$$\tau_n^2(\tilde{\sigma}_n^2) = \tfrac{1}{n}\sum_{i=1}^n \text{Var}(\epsilon_{n,i}^2) + \tfrac{4}{n^2}\sigma_0^4 \tau_n^2(\tilde{\lambda}_n)\text{tr}^2(H_n G_n) + \tfrac{4}{n^2}\text{Cov}(\epsilon_n'\epsilon, \epsilon_n' B_n \epsilon_n + c_n'\epsilon_n)\text{tr}(H_n G_n)\Phi_n^{-1},$$

where $\text{Cov}(\epsilon_n'\epsilon, \epsilon_n' B_n \epsilon_n + c_n'\epsilon_n)$ can be easily derived but not needed in light of Footnote 7.

Proof of Theorem 4: To prove the consistency of $\tilde{\tau}_n^2(\tilde{\lambda}_n)$ as an estimator of $\tau_n^2(\tilde{\lambda}_n)$, we need to prove (*a*) $\tilde{\Phi}_n - \Phi_n = o_p(1)$, and (*b*) $\tilde{\tau}_n^2(\tilde{\psi}_n^*) - \tau_n^2(\tilde{\psi}_n^*) = o_p(1)$. First, (*a*) follows from the proof of Theorem 2 (the asymptotic normality part). To prove (*b*), as $\tilde{\sigma}_n^2 = \sigma_0^2 + o_p(1)$ by Theorem 3, it suffices to show that, by the consistency of $\tilde{\theta}_n$ and referring to (18) and (19),

$$\tfrac{1}{n}\sum_{i=1}^n \left(\epsilon_{n,i}^2 \xi_{n,i}^2 - \text{Var}(\epsilon_{n,i}\xi_{n,i})\right) = o_p(1),$$

where $\xi_{n,i} = \zeta_{n,i} + b_{n,ii}\epsilon_{n,i} + c_{n,i}$. This follows immediately by Theorem A.1 and the poof of Theorem 1 of Baltagi and Yang (2013b).

The consistency of $\tilde{\tau}_n^2(\tilde{\beta}_n)$ follows that of $\tilde{\tau}_n^2(\tilde{\lambda}_n)$ and the consistency of $\tilde{\theta}_n$.

Finally, the same procedure proves the same set of the results for the regular QMLEs $\hat{\beta}_n$ and $\hat{\sigma}_n^2$.

References

Amemiya, T., 1985. *Advanced Econometrics.* Cambridge, Massachusetts: Harvard University Press.

Anselin, L., 1988. *Spatial Econometrics: Methods and Models.* The Netherlands: Kluwer Academic Publishers.

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

Anselin, L., Bera, A. K., 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: *Handbook of Applied Economic Statistics, Edited by Aman Ullah and David E. A. Giles.* New York: Marcel Dekker.

Arraiz, I., Drukker, D. M., Kelejian H. H., Prucha, I. R., 2010. A spatial Cliff-Ord type model with heteroskedastic innovations: small and large sample results. *Journal of Regional Science*, 50, 592-614.

Badinger, H., Egger, P., 2011. Estimation of higher-order spatial autoregressive cross-section models with heteroskedastic disturbances. *Papers in Regional Science* 90, 213-235.

Baltagi, B., Egger, P., Pfaffermayr, M., 2007. Estimating models of complex FDI: are there third country effects? *Journal of Econometrics* 140, 260-281.

Baltagi, B., Yang, Z. L., 2013a. Standardized LM tests for spatial error dependence in linear or panel regressions. *The Econometrics Journal* 16 103-134.

Baltagi, B., Yang, Z. L., 2013b. Heteroskedasticity and non-normality robust LM tests of spatial dependence. *Regional Science and Urban Economics* 43, 725-739.

Breusch, T., Pagan, A., 1979. A simple test for heteroskedasticity and random coefficient variation. *Econometrica* 47, 1287-1294.

Case, T., 1991. Spatial patterns in household demand. *Econometrica.* 59, 953-965.

Cliff, A., Ord, J. K., 1972. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* 4, 267-284.

Cliff, A. D., Ord, J. K., 1973. Spatial Autocorrelation. London: Pion.

Cliff, A. D., Ord, J. K., 1981. *Spatial Process, Models and Applications.* London: Pion.

Doğan, O., Taşpinar, S., 2014. Spatial autoregressive models with unknown heteroskedasticity: a comparison of Bayesian and robust GMM approach. *Regional Science and Urban Economics* 45, 1-21.

Glaeser, E. L., Sacerdote, B., Scheinkman, J. A., 1996. Crime and social interactions. *Quarterly Journal of Economics* 111, 507-548.

Hanushek, E. A., Kain, J. F., Markman, J. M., Rivkin, S. G., 2003. Does peer ability affect student achievement? *Journal of Applied Econometrics* 18, 527-544.

Jin, F., Lee, L. F., 2012. Approximated likelihood and root estimators for spatial interaction in spatial autoregressive models. *Regional Science and Urban Economics.* 42 446-458.

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

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

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

Kelejian, H.H., Prucha, I.R., 2007. HAC estimation in a spatial framework. *Journal of Econometrics* 140, 131-154.

Kelejian H. H., Prucha, I. R., 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics* 157 53-67.

Lee, L. F., 2001. Generalised method of moments estimation of spatial autoregressive processes - Unpublished manuscript.

Lee, L. F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* 72, 1899-1925.

Lee, L. F., Liu, X., 2010. Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. *Econometric Theory* 26, 187-230.

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

LeSage, J., 1997. Bayesian estimation of spatial autoregressive models. *International Regional Science Review* 20, 113-129.

LeSage, J., Pace, R. K., 2009. *Introduction to spatial econometrics.* New York: CRC Press.

Ord, J., 1975. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70, 120-126.

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

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

van der Vaart, A. W., 1998. *Asymptotic Statistics.* Cambridge University Press.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 817-838.

Yang, Z. L., 2010. A robust LM test for spatial error components. *Regional Science and Urban Economics* 40, 299-310.

Table 1: Empirical Mean(rmse)[sd]{$\hat{\text{sd}}$} of Estimators of λ for SAR Model
Cases when Regular QMLE is Consistent

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
DGP 1: Constant Circular Neighbours (REG-1), $\beta_0 = (3,1,1)'$					
.50	100	.464 (.105)[.098]{.092}	.473(.117)[.114]{.099}	.469(.121)[.117]	.479(.132)[.130]
	250	.488(.061)[.060]{.064}	.492(.063)[.063]{.059}	.489(.064)[.063]	.494(.071)[.071]
	500	.494(.043)[.043]{.046}	.497(.043)[.043]{.042}	.495(.043)[.043]	.498(.048)[.048]
	1000	.497(.030)[.030]{.032}	.498(.030)[.030]{.029}	.498(.030)[.030]	.498(.033)[.033]
.25	100	.212(.133)[.127]{.115}	.230(.128)[.127]{.123}	.221(.132)[.129]	.232(.146)[.145]
	250	.233(.080)[.078]{.078}	.246(.081)[.081]{.079}	.242(.082)[.081]	.247(.090)[.090]
	500	.245(.052)[.052]{.054}	.245(.054)[.054]{.054}	.243(.054)[.054]	.244(.060)[.059]
	1000	.246(.041)[.041]{.040}	.247(.039)[.039]{.038}	.246(.039)[.039]	.247(.043)[.043]
.00	100	-.033(.153)[.149]{.142}	-.014(.150)[.149]{.142}	-.024(.156)[.154]	-.009(.172)[.172]
	250	-.017 (.090)[.089]{.089}	-.007(.091)[.091]{.089}	-.011(.092)[.092]	-.005(.102)[.102]
	500	-.006 (.063)[.063]{.062}	-.002(.061)[.061]{.064}	-.004(.061)[.061]	-.002(.069)[.069]
	1000	-.006(.046)[.046]{.046}	-.003(.043)[.043]{.045}	-.005(.043)[.043]	-.003(.047)[.047]
-.25	100	-.285(.155)[.151]{.149}	-.272(.171)[.169]{.167}	-.286(.176)[.173]	-.275(.200)[.198]
	250	-.266(.101)[.100]{.100}	-.258(.100)[.100]{.099}	-.264(.101)[.100]	-.260(.112)[.112]
	500	-.259(.070)[.070]{.072}	-.255(.070)[.070]{.070}	-.258(.070)[.070]	-.256(.077)[.076]
	1000	-.253(.050)[.050]{.050}	-.250(.050)[.050]{.049}	-.252(.050)[.050]	-.250(.055)[.055]
-.50	100	-.524(.172)[.170]{.179}	-.506(.172)[.172]{.162}	-.521(.175)[.174]	-.513(.195)[.194]
	250	-.515(.108)[.107]{.112}	-.505(.104)[.104]{.101}	-.511(.104)[.104]	-.507(.117)[.116]
	500	-.501(.075)[.075]{.080}	-.497(.075)[.075]{.073}	-.501(.075)[.075]	-.497(.084)[.084]
	1000	-.500(.054)[.054]{.058}	-.499(.051)[.051]{.051}	-.500(.051)[.051]	-.500(.057)[.057]
DGP 2: Constant Circular Neighbours (REG-1), $\beta_0 = (3,1,1)'$					
.50	100	.465(.098)[.091]{.093}	.481(.107)[.105]{.099}	.475(.118)[.115]	.488(.142)[.141]
	250	.487(.062)[.061]{.063}	.494(.061)[.060]{.059}	.491(.061)[.061]	.495(.084)[.084]
	500	.494(.041)[.041]{.042}	.499(.042)[.042]{.040}	.497(.042)[.042]	.500(.059)[.059]
	1000	.498(.028)[.028]{.028}	.500(.028)[.028]{.029}	.499(.029)[.029]	.499(.041)[.041]
.25	100	.219(.129)[.126]{.124}	.238(.125)[.125]{.124}	.230(.128)[.127]	.251(.168)[.168]
	250	.236(.081)[.080]{.080}	.243(.080)[.079]{.079}	.239(.081)[.080]	.245(.108)[.108]
	500	.246(.056)[.056]{.059}	.250(.056)[.056]{.053}	.248(.056)[.056]	.251(.080)[.080]
	1000	.249(.039)[.039]{.041}	.251(.039)[.039]{.037}	.250(.039)[.039]	.250(.052)[.052]
.00	100	-.029(.146)[.143]{.139}	-.010(.143)[.143]{.139}	-.020(.150)[.148]	-.005(.209)[.209]
	250	-.011(.088)[.088]{.087}	-.003(.088)[.088]{.085}	-.008(.089)[.088]	.003(.122)[.122]
	500	-.005(.063)[.063]{.061}	-.008(.064)[.064]{.062}	-.010(.064)[.064]	-.004(.092)[.092]
	1000	-.003(.045)[.045]{.045}	-.001(.043)[.043]{.044}	-.003(.043)[.043]	.000(.060)[.060]
-.25	100	-.276(.158)[.155]{.145}	-.257(.156)[.156]{.153}	-.271(.160)[.159]	-.249(.223)[.223]
	250	-.268(.100)[.099]{.106}	-.261(.099)[.099]{.093}	-.266(.100)[.099]	-.260(.136)[.136]
	500	-.256(.073)[.073]{.077}	-.252(.073)[.073]{.069}	-.255(.074)[.073]	-.254(.102)[.102]
	1000	-.254(.050)[.050]{.050}	-.252(.049)[.049]{.048}	-.253(.050)[.049]	-.252(.068)[.068]
-.50	100	-.527(.155)[.153]{.163}	-.505(.154)[.154]{.154}	-.519(.158)[.157]	-.511(.221)[.221]
	250	-.505(.101)[.101]{.103}	-.500(.099)[.099]{.097}	-.506(.100)[.100]	-.502(.138)[.138]
	500	-.507(.075)[.075]{.077}	-.502(.072)[.072]{.072}	-.505(.072)[.072]	-.501(.103)[.103]
	1000	-.505(.050)[.049]{.049}	-.503(.050)[.049]{.050}	-.504(.050)[.050]	-.505(.071)[.071]

Table 1: Cont'd

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
DGP 3: Constant Circular Neighbours (REG-1), $\beta_0 = (3, 1, 1)'$					
.50	100	.474(.086)[.082]{.094}	.484(.096)[.095]{.089}	.476(.100)[.098]	.480(.149)[.148]
	250	.491(.057)[.056]{.054}	.497(.056)[.056]{.052}	.495(.076)[.076]	.499(.088)[.088]
	500	.493(.040)[.039]{.038}	.496(.040)[.039]{.038}	.494(.040)[.039]	.494(.067)[.067]
	1000	.496(.030)[.030]{.029}	.497(.029)[.028]{.027}	.497(.029)[.029]	.498(.045)[.045]
.25	100	.213(.124)[.119]{.110}	.231(.119)[.117]{.115}	.221(.125)[.122]	.233(.185)[.184]
	250	.240(.072)[.071]{.079}	.247(.071)[.070]{.067}	.242(.072)[.072]	.244(.116)[.116]
	500	.245(.050)[.050]{.052}	.247(.054)[.054]{.050}	.245(.055)[.054]	.245(.087)[.087]
	1000	.248(.037)[.037]{.038}	.250(.037)[.037]{.035}	.249(.037)[.037]	.250(.057)[.057]
.00	100	-.024(.124)[.122]{.116}	-.015(.140)[.140]{.143}	-.027(.148)[.145]	-.018(.221)[.220]
	250	-.010(.085)[.085]{.082}	-.002(.084)[.084]{.088}	-.007(.086)[.086]	-.002(.133)[.133]
	500	-.006(.059)[.058]{.060}	-.002(.058)[.058]{.058}	-.005(.059)[.059]	-.007(.101)[.101]
	1000	-.004(.045)[.044]{.044}	-.002(.042)[.042]{.041}	-.003(.043)[.043]	.000(.069)[.069]
-.25	100	-.276(.148)[.146]{.156}	-.258(.146)[.146]{.142}	-.272(.152)[.150]	-.261(.236)[.236]
	250	-.260(.093)[.092]{.101}	-.252(.093)[.093]{.096}	-.259(.094)[.093]	-.253(.153)[.153]
	500	-.256(.063)[.063]{.065}	-.254(.065)[.065]{.064}	-.256(.066)[.066]	-.251(.111)[.111]
	1000	-.254(.049)[.049]{.047}	-.250(.049)[.049]{.046}	-.252(.050)[.050]	-.251(.076)[.076]
-.50	100	-.514(.141)[.140]{.153}	-.508(.161)[.161]{.167}	-.526(.165)[.163]	-.513(.246)[.245]
	250	-.511(.092)[.091]{.098}	-.506(.097)[.097]{.091}	-.512(.099)[.098]	-.514(.155)[.154]
	500	-.503(.069)[.069]{.069}	-.499(.069)[.069]{.067}	-.503(.069)[.069]	-.498(.111)[.111]
	1000	-.503(.051)[.051]{.051}	-.501(.051)[.051]{.049}	-.503(.051)[.051]	-.505(.081)[.081]
DGP 1: Queen Contiguity (REG-1), $\beta_0 = (.3, .1, .1)'$					
.50	100	.447(.156)[.146]{.136}	.471(.147)[.144]{.148}	.463(.158)[.154]	.501(.207)[.207]
	250	.482(.081)[.079]{.088}	.495(.079)[.079]{.079}	.488(.081)[.080]	.499(.085)[.085]
	500	.489(.061)[.059]{.063}	.494(.056)[.056]{.056}	.491(.070)[.069]	.497(.071)[.071]
	1000	.496(.041)[.041]{.045}	.497(.042)[.042]{.040}	.495(.042)[.042]	.498(.043)[.043]
.25	100	.207(.170)[.165]{.155}	.231(.167)[.166]{.155}	.219(.172)[.169]	.240(.186)[.186]
	250	.232(.103)[.101]{.101}	.241(.102)[.102]{.099}	.234(.104)[.102]	.242(.106)[.106]
	500	.242(.072)[.072]{.072}	.249(.072)[.072]{.070}	.245(.072)[.072]	.250(.074)[.074]
	1000	.244(.050)[.050]{.052}	.247(.050)[.050]{.050}	.245(.050)[.050]	.247(.051)[.051]
.00	100	-.046(.192)[.186]{.173}	-.021(.188)[.187]{.174}	-.036(.195)[.192]	-.021(.205)[.204]
	250	-.019(.117)[.115]{.112}	-.008(.115)[.115]{.112}	-.017(.117)[.116]	-.010(.120)[.120]
	500	-.008(.080)[.080]{.079}	-.001(.080)[.080]{.080}	-.005(.080)[.080]	-.001(.082)[.082]
	1000	-.005(.058)[.058]{.057}	-.002(.058)[.058]{.057}	-.004(.058)[.058]	-.002(.059)[.059]
-.25	100	-.286(.199)[.195]{.192}	-.258(.198)[.198]{.193}	-.277(.205)[.204]	-.264(.218)[.217]
	250	-.272(.122)[.120]{.125}	-.258(.121)[.120]{.120}	-.268(.122)[.121]	-.265(.126)[.125]
	500	-.260(.089)[.088]{.089}	-.253(.089)[.089]{.086}	-.258(.089)[.089]	-.256(.090)[.090]
	1000	-.256(.063)[.063]{.064}	-.252(.063)[.063]{.061}	-.255(.063)[.063]	-.254(.064)[.064]
-.50	100	-.526(.194)[.192]{.201}	-.502(.194)[.194]{.187}	-.521(.197)[.196]	-.521(.214)[.213]
	250	-.513(.122)[.121]{.128}	-.501(.122)[.122]{.122}	-.513(.124)[.123]	-.514(.128)[.127]
	500	-.504(.087)[.087]{.088}	-.498(.088)[.088]{.087}	-.503(.088)[.088]	-.503(.089)[.089]
	1000	-.503(.063)[.063]{.061}	-.500(.063)[.063]{.063}	-.502(.063)[.063]	-.502(.064)[.064]

Table 1: Cont'd

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
		DGP 2: Queen Contiguity (REG-1), $\beta_0 = (.3, .1, .1)'$			
.50	100	.455(.136)[.129]{.137}	.481(.129)[.128]{.123}	.470(.135)[.132]	.581(.354)[.345]
	250	.480(.087)[.083]{.100}	.493(.078)[.078]{.076}	.487(.080)[.079]	.533(.160)[.157]
	500	.490(.057)[.056]{.057}	.497(.056)[.056]{.054}	.495(.068)[.068]	.518(.088)[.086]
	1000	.496(.042)[.042]{.047}	.499(.042)[.042]{.039}	.498(.042)[.042]	.510(.053)[.052]
.25	100	.206(.171)[.166]{.155}	.233(.166)[.165]{.161}	.224(.180)[.178]	.308(.366)[.361]
	250	.222(.108)[.104]{.105}	.240(.097)[.096]{.094}	.232(.099)[.098]	.272(.139)[.137]
	500	.239(.072)[.071]{.076}	.246(.071)[.071]{.068}	.242(.072)[.071]	.259(.089)[.089]
	1000	.246(.050)[.050]{.050}	.245(.052)[.052]{.050}	.244(.053)[.052]	.257(.070)[.070]
.00	100	-.035(.177)[.174]{.165}	-.023(.184)[.182]{.188}	-.039(.191)[.187]	.002(.243)[.243]
	250	-.019(.116)[.115]{.109}	-.005(.115)[.114]{.106}	-.014(.117)[.116]	.016(.153)[.152]
	500	-.009(.081)[.080]{.078}	-.004(.081)[.081]{.077}	-.008(.082)[.081]	.012(.105)[.105]
	1000	-.004(.057)[.057]{.057}	-.002(.057)[.057]{.056}	-.005(.057)[.057]	.007(.069)[.069]
-.25	100	-.283(.185)[.182]{.190}	-.268(.186)[.185]{.186}	-.285(.192)[.189]	-.254(.251)[.251]
	250	-.270(.122)[.120]{.125}	-.256(.121)[.120]{.114}	-.267(.123)[.122]	-.253(.161)[.161]
	500	-.256(.085)[.084]{.085}	-.250(.085)[.085]{.082}	-.254(.085)[.085]	-.242(.106)[.106]
	1000	-.252(.063)[.063]{.060}	-.249(.063)[.063]{.060}	-.251(.063)[.063]	-.245(.078)[.078]
-.50	100	-.518(.195)[.194]{.204}	-.506(.188)[.187]{.180}	-.529(.193)[.190]	-.523(.255)[.254]
	250	-.513(.127)[.126]{.128}	-.501(.127)[.127]{.125}	-.512(.128)[.128]	-.513(.168)[.167]
	500	-.505(.088)[.088]{.084}	-.500(.089)[.089]{.085}	-.505(.089)[.088]	-.500(.110)[.110]
	1000	-.503(.063)[.063]{.060}	-.500(.063)[.063]{.061}	-.503(.063)[.063]	-.501(.077)[.077]
		DGP 3: Queen Contiguity (REG-1), $\beta_0 = (.3, .1, .1)'$			
.50	100	.453(.128)[.119]{.126}	.479(.120)[.118]{.109}	.470(.144)[.141]	.631(.463)[.444]
	250	.479(.079)[.076]{.072}	.492(.076)[.075]{.069}	.487(.079)[.077]	.583(.287)[.275]
	500	.486(.056)[.054]{.057}	.492(.054)[.054]{.049}	.489(.055)[.054]	.554(.206)[.198]
	1000	.494(.039)[.038]{.031}	.497(.039)[.038]{.037}	.496(.039)[.039]	.530(.107)[.103]
.25	100	.205(.151)[.144]{.146}	.232(.145)[.144]{.148}	.220(.154)[.151]	.354(.469)[.458]
	250	.231(.100)[.098]{.100}	.245(.098)[.098]{.095}	.237(.100)[.099]	.307(.277)[.271]
	500	.237(.071)[.070]{.072}	.244(.070)[.070]{.069}	.240(.071)[.070]	.306(.250)[.244]
	1000	.246(.049)[.049]{.055}	.248(.050)[.050]{.049}	.246(.051)[.050]	.271(.126)[.124]
.00	100	-.048(.164)[.157]{.159}	-.015(.169)[.168]{.164}	-.029(.175)[.172]	.057(.327)[.321]
	250	-.018(.106)[.104]{.104}	-.004(.104)[.104]{.099}	-.013(.107)[.106]	.038(.214)[.210]
	500	-.011(.077)[.076]{.075}	-.003(.077)[.076]{.071}	-.008(.077)[.077]	.032(.169)[.166]
	1000	-.004(.055)[.055]{.055}	-.001(.055)[.055]{.053}	-.003(.055)[.055]	.028(.132)[.129]
-.25	100	-.284(.170)[.167]{.179}	-.263(.169)[.168]{.163}	-.284(.175)[.172]	-.245(.283)[.283]
	250	-.268(.119)[.117]{.110}	-.254(.118)[.117]{.115}	-.265(.120)[.119]	-.220(.214)[.211]
	500	-.258(.081)[.081]{.083}	-.252(.081)[.081]{.079}	-.257(.081)[.081]	-.221(.176)[.174]
	1000	-.252(.059)[.059]{.054}	-.254(.059)[.059]{.056}	-.256(.059)[.059]	-.224(.151)[.148]
-.50	100	-.523(.176)[.175]{.189}	-.516(.182)[.182]{.187}	-.539(.192)[.188]	-.528(.312)[.311]
	250	-.514(.120)[.119]{.113}	-.501(.119)[.119]{.118}	-.513(.120)[.119]	-.501(.215)[.215]
	500	-.503(.085)[.085]{.084}	-.500(.085)[.085]{.088}	-.505(.085)[.085]	-.491(.172)[.172]
	1000	-.503(.063)[.063]{.061}	-.500(.063)[.063]{.059}	-.502(.063)[.063]	-.496(.150)[.150]

Table 2: Empirical Mean(rmse)[sd]{$\hat{\text{sd}}$} of Estimators of λ for SAR Model
Case I of Inconsistent QMLE: Circular Neighbours (REG-1)

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
DGP 1: $\beta_0 = (3, 1, 1)'$					
.50	100	.434(.119)[.100]	.481(.103)[.101]{.093}	.477(.107)[.104]	.483(.113)[.112]
	250	.458(.071)[.057]	.491(.059)[.059]{.057}	.489(.058)[.057]	.492(.061)[.060]
	500	.463(.056)[.043]	.496(.044)[.044]{.043}	.495(.043)[.043]	.496(.046)[.046]
	1000	.472(.040)[.028]	.500(.029)[.029]{.028}	.499(.028)[.028]	.500(.030)[.030]
.25	100	.197(.120)[.107]	.233(.116)[.115]{.115}	.226(.117)[.115]	.232(.127)[.125]
	250	.218(.077)[.070]	.242(.075)[.074]{.070}	.239(.073)[.072]	.242(.075)[.075]
	500	.222(.060)[.053]	.246(.057)[.057]{.054}	.245(.057)[.057]	.247(.061)[.060]
	1000	.225(.042)[.034]	.246(.037)[.036]{.035}	.245(.036)[.036]	.246(.038)[.038]
.00	100	-.023(.114)[.111]	-.009(.127)[.126]{.127}	-.015(.127)[.127]	-.006(.136)[.136]
	250	-.012(.073)[.072]	-.007(.081)[.080]{.078}	-.009(.080)[.079]	-.005(.084)[.084]
	500	-.005(.054)[.053]	-.002(.060)[.060]{.060}	-.003(.060)[.060]	-.001(.064)[.064]
	1000	-.002(.036)[.036]	-.001(.040)[.040]{.039}	-.002(.039)[.039]	-.001(.042)[.042]
-.25	100	-.249(.110)[.110]	-.271(.137)[.135]{.139}	-.271(.132)[.131]	-.270(.155)[.154]
	250	-.226(.072)[.068]	-.250(.082)[.081]{.080}	-.251(.076)[.076]	-.250(.081)[.081]
	500	-.224(.058)[.052]	-.252(.063)[.063]{.062}	-.252(.060)[.060]	-.251(.064)[.064]
	1000	-.225(.043)[.034]	-.252(.040)[.040]{.040}	-.252(.039)[.039]	-.252(.042)[.042]
-.50	100	-.449(.105)[.092]	-.494(.114)[.114]{.119}	-.492(.105)[.104]	-.498(.112)[.112]
	250	-.448(.079)[.059]	-.503(.076)[.076]{.076}	-.498(.065)[.065]	-.500(.070)[.070]
	500	-.444(.073)[.046]	-.506(.061)[.061]{.059}	-.505(.054)[.054]	-.506(.057)[.056]
	1000	-.444(.064)[.030]	-.501(.037)[.037]{.037}	-.500(.034)[.034]	-.501(.035)[.035]
DGP 2: $\beta_0 = (3, 1, 1)'$					
.50	100	.438(.114)[.096]	.483(.098)[.097]{.089}	.477(.105)[.102]	.485(.130)[.129]
	250	.462(.066)[.054]	.495(.055)[.055]{.055}	.492(.053)[.053]	.496(.067)[.067]
	500	.467(.054)[.043]	.500(.044)[.044]{.042}	.498(.043)[.043]	.499(.057)[.057]
	1000	.473(.039)[.027]	.501(.028)[.028]{.028}	.500(.027)[.027]	.501(.034)[.034]
.25	100	.201(.123)[.113]	.236(.120)[.119]{.109}	.228(.122)[.120]	.235(.147)[.146]
	250	.219(.072)[.066]	.244(.070)[.070]{.069}	.242(.070)[.069]	.245(.087)[.087]
	500	.220(.059)[.051]	.244(.055)[.054]{.053}	.243(.054)[.054]	.247(.071)[.071]
	1000	.228(.040)[.033]	.248(.035)[.035]{.035}	.248(.035)[.034]	.249(.043)[.043]
.00	100	-.022(.116)[.114]	-.010(.131)[.131]{.129}	-.016(.129)[.128]	-.005(.159)[.158]
	250	-.010(.073)[.072]	-.005(.081)[.081]{.079}	-.008(.080)[.079]	-.004(.097)[.096]
	500	-.004(.051)[.051]	-.001(.058)[.058]{.058}	-.002(.057)[.057]	.001(.075)[.075]
	1000	-.003(.036)[.036]	-.002(.040)[.040]{.039}	-.002(.039)[.039]	-.001(.048)[.048]
-.25	100	-.239(.109)[.108]	-.257(.131)[.131]{.129}	-.256(.122)[.122]	-.248(.150)[.150]
	250	-.232(.071)[.069]	-.257(.083)[.082]{.079}	-.257(.077)[.077]	-.253(.093)[.093]
	500	-.223(.059)[.052]	-.251(.062)[.062]{.060}	-.251(.060)[.060]	-.247(.078)[.078]
	1000	-.222(.045)[.036]	-.249(.041)[.041]{.040}	-.249(.040)[.040]	-.249(.048)[.048]
-.50	100	-.452(.105)[.093]	-.499(.114)[.114]{.116}	-.495(.110)[.110]	-.496(.123)[.123]
	250	-.448(.080)[.061]	-.501(.073)[.073]{.073}	-.499(.066)[.066]	-.499(.079)[.079]
	500	-.438(.077)[.046]	-.500(.059)[.059]{.058}	-.498(.052)[.052]	-.497(.065)[.065]
	1000	-.444(.064)[.031]	-.501(.037)[.037]{.037}	-.502(.034)[.034]	-.502(.041)[.041]

Table 2: Cont'd

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
			DGP 3: $\beta_0 = (3, 1, 1)'$		
.50	100	.445(.107)[.092]	.486(.087)[.086]{.079}	.482(.092)[.090]	.493(.144)[.144]
	250	.464(.066)[.055]	.495(.054)[.054]{.049}	.493(.054)[.053]	.497(.073)[.073]
	500	.467(.055)[.044]	.497(.041)[.041]{.039}	.496(.042)[.041]	.497(.060)[.060]
	1000	.473(.040)[.030]	.499(.027)[.027]{.026}	.499(.027)[.027]	.500(.037)[.037]
.25	100	.199(.116)[.105]	.230(.110)[.108]{.099}	.241(.068)[.067]	.245(.090)[.089]
	250	.219(.071)[.064]	.243(.069)[.068]{.063}	.241(.068)[.067]	.245(.090)[.089]
	500	.222(.058)[.050]	.244(.054)[.053]{.049}	.243(.053)[.053]	.242(.078)[.078]
	1000	.228(.040)[.033]	.248(.035)[.034]{.033}	.248(.034)[.034]	.250(.045)[.045]
.00	100	-.019(.107)[.105]	-.008(.120)[.120]{.119}	-.013(.119)[.119]	-.005(.164)[.164]
	250	-.008(.065)[.065]	-.003(.072)[.072]{.069}	-.006(.072)[.072]	-.003(.101)[.101]
	500	-.006(.051)[.050]	-.004(.057)[.057]{.054}	-.006(.058)[.058]	-.007(.089)[.089]
	1000	-.003(.035)[.034]	-.002(.038)[.038]{.037}	-.003(.038)[.038]	-.003(.053)[.053]
-.25	100	-.243(.102)[.102]	-.260(.123)[.123]{.120}	-.262(.118)[.117]	-.257(.157)[.156]
	250	-.230(.072)[.069]	-.250(.077)[.077]{.072}	-.251(.074)[.074]	-.248(.098)[.098]
	500	-.228(.055)[.050]	-.255(.058)[.058]{.056}	-.256(.058)[.057]	-.255(.083)[.083]
	1000	-.223(.044)[.035]	-.250(.039)[.039]{.038}	-.250(.039)[.039]	-.249(.052)[.052]
-.50	100	-.450(.107)[.095]	-.486(.110)[.109]{.112}	-.485(.105)[.104]	-.484(.125)[.123]
	250	-.450(.081)[.063]	-.502(.074)[.074]{.070}	-.498(.064)[.064]	-.496(.085)[.085]
	500	-.439(.081)[.053]	-.499(.061)[.061]{.059}	-.497(.051)[.051]	-.499(.069)[.069]
	1000	-.445(.066)[.037]	-.500(.038)[.038]{.036}	-.500(.034)[.034]	-.501(.044)[.044]
			DGP 1: $\beta_0 = (.3, .1, .1)'$		
.50	100	.407(.154)[.123]	.474(.129)[.127]{.119}	.467(.148)[.144]	.499(.189)[.189]
	250	.437(.100)[.078]	.489(.080)[.079]{.075}	.485(.082)[.080]	.494(.083)[.083]
	500	.445(.076)[.053]	.494(.054)[.054]{.053}	.493(.069)[.069]	.497(.066)[.066]
	1000	.453(.060)[.037]	.499(.037)[.037]{.038}	.498(.037)[.037]	.500(.038)[.038]
.25	100	.174(.156)[.136]	.226(.155)[.153]{.149}	.213(.165)[.161]	.235(.195)[.194]
	250	.199(.101)[.087]	.238(.097)[.097]{.096}	.233(.100)[.098]	.241(.102)[.102]
	500	.208(.076)[.063]	.243(.069)[.069]{.068}	.240(.070)[.069]	.243(.070)[.070]
	1000	.213(.058)[.045]	.246(.049)[.049]{.048}	.245(.050)[.049]	.246(.050)[.050]
.00	100	-.041(.146)[.140]	-.023(.170)[.168]{.165}	-.040(.179)[.174]	-.026(.184)[.182]
	250	-.016(.096)[.095]	-.009(.114)[.113]{.117}	-.015(.115)[.114]	-.009(.116)[.116]
	500	-.008(.066)[.066]	-.004(.078)[.078]{.077}	-.008(.079)[.079]	-.005(.080)[.080]
	1000	-.004(.044)[.044]	-.002(.052)[.052]{.054}	-.003(.052)[.052]	-.002(.053)[.053]
-.25	100	-.240(.136)[.136]	-.270(.176)[.175]{.172}	-.292(.185)[.180]	-.291(.201)[.197]
	250	-.213(.095)[.087]	-.251(.110)[.110]{.111}	-.259(.111)[.111]	-.256(.114)[.114]
	500	-.210(.074)[.062]	-.252(.079)[.079]{.079}	-.256(.079)[.079]	-.255(.080)[080]
	1000	-.209(.060)[.044]	-.252(.055)[.055]{.056}	-.254(.055)[.055]	-.254(.056)[.056]
-.50	100	-.417(.149)[.124]	-.496(.164)[.164]{.159}	-.531(.202)[.199]	-.535(.213)[.210]
	250	-.413(.117)[.078]	-.504(.103)[.103]{.102}	-.512(.103)[.102]	-.516(.107)[.106]
	500	-.409(.107)[.056]	-.501(.073)[.073]{.073}	-.506(.073)[.073]	-.507(.074)[.074]
	1000	-.405(.103)[.039]	-.498(.051)[.051]{.052}	-.501(.051)[.051]	-.501(.051)[.051]

Table 2: Cont'd

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
		DGP 2: $\beta_0 = (.3, .1, .1)'$			
.50	100	.416(.147)[.121]	.482(.123)[.121]{.119}	.475(.138)[.136]	.592(.342)[.329]
	250	.438(.101)[.080]	.490(.081)[.080]{.079}	.487(.090)[.089]	.528(.157)[.154]
	500	.448(.074)[.053]	.496(.053)[.053]{.052}	.494(.054)[.053]	.511(.068)[.067]
	1000	.452(.061)[.038]	.499(.038)[.038]{.037}	.498(.038)[.038]	.508(.047)[.047]
.25	100	.184(.152)[.137]	.236(.154)[.154]{.157}	.224(.165)[.163]	.304(.305)[.301]
	250	.203(.100)[.088]	.242(.097)[.097]{.091}	.236(.099)[.098]	.271(.149)[.147]
	500	.211(.073)[.062]	.246(.067)[.067]{.066}	.243(.068)[.068]	.264(.109)[.109]
	1000	.217(.055)[.044]	.250(.048)[.048]{.047}	.249(.048)[.048]	.258(.058)[.058]
.00	100	-.040(.144)[.139]	-.021(.171)[.169]{.164}	-.039(.180)[.176]	.014(.262)[.262]
	250	-.016(.091)[.089]	-.010(.107)[.107]{.104}	-.016(.109)[.108]	.008(.134)[.134]
	500	-.007(.063)[.063]	-.003(.075)[.075]{.074}	-.006(.075)[.075]	.008(.090)[.090]
	1000	-.003(.046)[.046]	-.001(.054)[.054]{.053}	-.003(.054)[.054]	.006(.066)[.066]
-.25	100	-.232(.133)[.131]	-.259(.169)[.169]{.159}	-.281(.180)[.177]	-.254(.266)[.266]
	250	-.216(.090)[.083]	-.254(.106)[.106]{.107}	-.262(.108)[.107]	-.249(.138)[.138]
	500	-.210(.073)[.061]	-.251(.077)[.077]{.077}	-.255(.077)[.077]	-.246(.088)[.088]
	1000	-.207(.063)[.046]	-.249(.057)[.057]{.055}	-.251(.057)[.057]	-.247(.067)[.067]
-.50	100	-.424(.148)[.127]	-.503(.163)[.163]{.160}	-.535(.191)[.187]	-.549(.246)[.241]
	250	-.410(.123)[.084]	-.499(.105)[.105]{.099}	-.507(.106)[.105]	-.513(.151)[.151]
	500	-.409(.108)[.058]	-.500(.071)[.071]{.072}	-.504(.071)[.071]	-.507(.086)[.086]
	1000	-.409(.100)[.041]	-.503(.050)[.050]{.051}	-.506(.051)[.050]	-.509(.063)[.062]
		DGP 3: $\beta_0 = (.3, .1, .1)'$			
.50	100	.416(.147)[.120]	.480(.118)[.116]{.099}	.473(.130)[.128]	.652(.453)[.426]
	250	.439(.096)[.074]	.490(.071)[.070]{.065}	.486(.073)[.071]	.572(.247)[.236]
	500	.449(.074)[.054]	.497(.050)[.050]{.048}	.495(.051)[.051]	.547(.189)[.184]
	1000	.453(.060)[.037]	.498(.034)[.034]{.035}	.497(.035)[.034]	.523(.104)[.101]
.25	100	.174(.153)[.133]	.224(.147)[.144]{.137}	.212(.156)[.152]	.335(.387)[.378]
	250	.210(.089)[.080]	.249(.087)[.087]{.083}	.243(.087)[.087]	.310(.245)[.237]
	500	.211(.072)[.061]	.244(.065)[.065]{.061}	.242(.066)[.065]	.283(.198)[.195]
	1000	.214(.057)[.044]	.247(.046)[.046]{.044}	.246(.047)[.046]	.266(.116)[.115]
.00	100	-.027(.135)[.133]	-.008(.161)[.160]{.153}	-.026(.172)[.170]	.077(.422)[.414]
	250	-.014(.087)[.086]	-.006(.103)[.103]{.099}	-.013(.105)[.104]	.052(.263)[.258]
	500	-.008(.059)[.058]	-.004(.070)[.070]{.069}	-.008(.071)[.070]	.026(.151)[.149]
	1000	-.003(.042)[.042]	-.001(.050)[.050]{.050}	-.003(.050)[.050]	.025(.116)[.114]
-.25	100	-.234(.131)[.130]	-.262(.172)[.172]{.179}	-.288(.184)[.180]	-.238(.295)[.295]
	250	-.218(.090)[.084]	-.254(.105)[.105]{.099}	-.262(.107)[.106]	-.223(.239)[.238]
	500	-.213(.073)[.063]	-.252(.076)[.076]{.071}	-.256(.077)[.076]	-.233(.161)[.160]
	1000	-.208(.062)[.046]	-.250(.055)[.055]{.053}	-.252(.055)[.055]	-.238(.128)[.127]
-.50	100	-.418(.151)[.127]	-.495(.158)[.158]{.151}	-.526(.178)[.176]	-.544(.304)[.301]
	250	-.411(.126)[.089]	-.503(.105)[.105]{.099}	-.511(.105)[.104]	-.508(.199)[.198]
	500	-.408(.113)[.066]	-.500(.073)[.073]{.069}	-.504(.072)[.072]	-.501(.156)[.156]
	1000	-.403(.109)[.049]	-.496(.051)[.051]{.049}	-.498(.051)[.051]	-.502(.129)[.129]

Table 3: Empirical Mean(rmse)[sd]{$\hat{\text{sd}}$} of Estimators of λ for SAR Model
Case II of Inconsistent QMLE: Group Interaction (REG-2)

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
			DGP 1: $\beta_0 = (3, 1, 1)'$		
.50	100	.422(.124)[.096]	.478(.102)[.099]{.093}	.469(.109)[.105]	.470(.112)[.108]
	250	.461(.069)[.057]	.493(.059)[.059]{.056}	.488(.061)[.060]	.491(.065)[.064]
	500	.472(.047)[.037]	.497(.039)[.038]{.038}	.494(.039)[.039]	.496(.041)[.041]
	1000	.476(.037)[.028]	.499(.029)[.029]{.028}	.497(.029)[.029]	.498(.031)[.030]
.25	100	.159(.161)[.132]	.224(.142)[.140]{.139}	.210(.156)[.150]	.215(.162)[.158]
	250	.210(.087)[.078]	.244(.082)[.081]{.080}	.237(.085)[.084]	.242(.090)[.090]
	500	.223(.060)[.053]	.247(.056)[.056]{.055}	.243(.057)[.057]	.246(.061)[.061]
	1000	.232(.042)[.037]	.251(.039)[.039]{.040}	.249(.040)[.040]	.251(.043)[.043]
.00	100	-.079(.179)[.160]	-.023(.183)[.181]{.183}	-.035(.194)[.191]	-.026(.203)[.201]
	250	-.034(.100)[.094]	-.011(.103)[.103]{.102}	-.020(.107)[.105]	-.014(.112)[.111]
	500	-.018(.067)[.065]	-.006(.071)[.070]{.070}	-.013(.072)[.071]	-.009(.075)[.075]
	1000	-.011(.049)[.048]	-.005(.052)[.052]{.051}	-.009(.054)[.053]	-.007(.057)[.057]
-.25	100	-.317(.184)[.171]	-.285(.210)[.207]{.213}	-.300(.222)[.216]	-.291(.234)[.231]
	250	-.264(.109)[.108]	-.266(.126)[.124]{.123}	-.276(.128)[.125]	-.271(.134)[.132]
	500	-.247(.074)[.074]	-.258(.085)[.085]{.084}	-.265(.086)[.085]	-.262(.091)[.090]
	1000	-.235(.056)[.054]	-.254(.061)[.060]{.060}	-.257(.062)[.061]	-.255(.065)[.065]
-.50	100	-.532(.181)[.178]	-.534(.226)[.224]{.219}	-.546(.231)[.226]	-.543(.245)[.241]
	250	-.468(.120)[.116]	-.505(.146)[.146]{.144}	-.515(.143)[.142]	-.511(.151)[.150]
	500	-.460(.090)[.080]	-.507(.101)[.100]{.097}	-.511(.096)[.095]	-.509(.101)[.101]
	1000	-.448(.078)[.057]	-.501(.070)[.070]{.069}	-.505(.069)[.069]	-.503(.073)[.073]
			DGP 2: $\beta_0 = (3, 1, 1)'$		
.50	100	.437(.117)[.098]	.492(.099)[.098]{.089}	.487(.110)[.110]	.497(.126)[.126]
	250	.465(.066)[.056]	.499(.057)[.057]{.054}	.494(.060)[.059]	.504(.074)[.074]
	500	.471(.047)[.037]	.497(.038)[.038]{.038}	.494(.039)[.038]	.499(.050)[.050]
	1000	.477(.035)[.027]	.500(.028)[.028]{.028}	.498(.028)[.028]	.500(.036)[.036]
.25	100	.167(.155)[.130]	.230(.137)[.135]{.129}	.220(.151)[.148]	.235(.172)[.171]
	250	.211(.085)[.076]	.245(.079)[.079]{.077}	.236(.082)[.081]	.245(.100)[.100]
	500	.219(.060)[.051]	.243(.054)[.053]{.054}	.238(.055)[.054]	.245(.067)[.067]
	1000	.231(.042)[.038]	.251(.040)[.040]{.039}	.248(.040)[.040]	.250(.052)[.052]
.00	100	-.084(.181)[.160]	-.028(.179)[.176]{.169}	-.044(.195)[.190]	-.019(.228)[.227]
	250	-.031(.098)[.093]	-.008(.101)[.101]{.098}	-.018(.107)[.105]	-.005(.134)[.134]
	500	-.015(.068)[.067]	-.003(.073)[.073]{.069}	-.009(.074)[.074]	.001(.095)[.095]
	1000	-.008(.050)[.049]	-.002(.053)[.053]{.050}	-.005(.054)[.054]	.000(.069)[.069]
-.25	100	-.313(.178)[.167]	-.283(.206)[.203]{.211}	-.296(.215)[.210]	-.268(.259)[.258]
	250	-.262(.109)[.108]	-.263(.126)[.126]{.119}	-.272(.128)[.126]	-.256(.159)[.159]
	500	-.243(.072)[.072]	-.254(.082)[.082]{.082}	-.260(.081)[.081]	-.252(.101)[.101]
	1000	-.235(.055)[.053]	-.253(.060)[.060]{.060}	-.256(.061)[.061]	-.252(.080)[.080]
-.50	100	-.523(.182)[.181]	-.531(.241)[.239]{.230}	-.541(.237)[.233]	-.510(.284)[.283]
	250	-.471(.118)[.114]	-.510(.142)[.142]{.140}	-.517(.138)[.137]	-.497(.174)[.174]
	500	-.458(.092)[.082]	-.503(.101)[.101]{.095}	-.509(.098)[.097]	-.498(.121)[.121]
	1000	-.445(.079)[.057]	-.497(.068)[.068]{.069}	-.500(.068)[.068]	-.493(.090)[.089]

Table 3: Cont'd

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
DGP 3: $\beta_0 = (3, 1, 1)'$					
.50	100	.433(.115)[.094]	.484(.090)[.089]{.081}	.476(.110)[.107]	.485(.138)[.138]
	250	.469(.062)[.054]	.500(.053)[.053]{.050}	.495(.055)[.055]	.503(.076)[.076]
	500	.473(.046)[.037]	.497(.036)[.036]{.035}	.494(.037)[.037]	.496(.051)[.051]
	1000	.478(.035)[.027]	.500(.026)[.026]{.026}	.498(.027)[.027]	.502(.038)[.038]
.25	100	.173(.145)[.123]	.232(.125)[.124]{.114}	.221(.150)[.147]	.236(.187)[.186]
	250	.211(.086)[.077]	.243(.079)[.079]{.071}	.236(.084)[.083]	.247(.115)[.115]
	500	.225(.056)[.051]	.248(.052)[.052]{.051}	.244(.054)[.053]	.250(.078)[.078]
	1000	.228(.044)[.038]	.246(.039)[.039]{.038}	.244(.040)[.039]	.248(.056)[.056]
.00	100	-.078(.169)[.150]	-.026(.174)[.172]{.164}	-.044(.188)[.183]	-.019(.229)[.228]
	250	-.030(.098)[.093]	-.008(.102)[.102]{.099}	-.018(.107)[.106]	-.002(.145)[.145]
	500	-.017(.066)[.063]	-.005(.069)[.069]{.066}	-.012(.071)[.070]	-.005(.097)[.097]
	1000	-.007(.047)[.046]	-.001(.050)[.050]{.048}	-.005(.051)[.051]	-.003(.073)[.073]
-.25	100	-.305(.178)[.170]	-.270(.197)[.196]{.199}	-.291(.218)[.214]	-.262(.280)[.280]
	250	-.262(.104)[.103]	-.264(.123)[.122]{.119}	-.272(.124)[.122]	-.256(.173)[.173]
	500	-.248(.071)[.071]	-.259(.081)[.080]{.078}	-.265(.082)[.081]	-.256(.115)[.115]
	1000	-.234(.055)[.053]	-.251(.059)[.059]{.057}	-.255(.060)[.060]	-.249(.090)[.090]
-.50	100	-.535(.181)[.177]	-.530(.218)[.216]{.223}	-.555(.236)[.229]	-.528(.304)[.303]
	250	-.474(.118)[.115]	-.515(.148)[.147]{.139}	-.523(.142)[.141]	-.505(.195)[.195]
	500	-.457(.091)[.080]	-.504(.094)[.093]{.092}	-.509(.091)[.090]	-.500(.125)[.125]
	1000	-.449(.081)[.063]	-.502(.069)[.069]{.067}	-.505(.069)[.069]	-.498(.101)[.101]
DGP 1: $\beta_0 = (.3, .1, .1)'$					
.50	100	.364(.203)[.150]	.456(.148)[.141]{.129}	.419(.219)[.204]	.423(.234)[.220]
	250	.433(.105)[.080]	.487(.079)[.078]{.073}	.468(.095)[.090]	.469(.095)[.090]
	500	.450(.073)[.053]	.494(.053)[.053]{.051}	.482(.057)[.054]	.483(.057)[.054]
	1000	.460(.054)[.036]	.497(.036)[.036]{.036}	.491(.038)[.037]	.491(.038)[.037]
.25	100	.092(.246)[.188]	.193(.206)[.197]{.185}	.126(.269)[.239]	.127(.289)[.261]
	250	.178(.129)[.107]	.232(.114)[.112]{.109}	.203(.126)[.116]	.202(.127)[.117]
	500	.202(.084)[.069]	.242(.074)[.073]{.073}	.225(.079)[.075]	.225(.079)[.075]
	1000	.215(.059)[.048]	.246(.051)[.051]{.051}	.238(.053)[.051]	.238(.053)[.051]
.00	100	-.150(.258)[.211]	-.070(.257)[.247]{.233}	-.161(.331)[.289]	-.159(.346)[.307]
	250	-.060(.141)[.127]	-.028(.148)[.146]{.133}	-.066(.164)[.150]	-.066(.165)[.151]
	500	-.030(.090)[.085]	-.011(.097)[.097]{.093}	-.033(.104)[.099]	-.032(.104)[.099]
	1000	-.016(.059)[.057]	-.007(.065)[.065]{.066}	-.018(.068)[.066]	-.018(.069)[.066]
-.25	100	-.365(.241)[.212]	-.328(.294)[.283]{.272}	-.441(.381)[.330]	-.432(.409)[.366]
	250	-.260(.127)[.126]	-.264(.159)[.158]{.156}	-.308(.172)[.162]	-.309(.173)[.162]
	500	-.243(.093)[.093]	-.263(.116)[.115]{.110}	-.289(.123)[.117]	-.289(.123)[.117]
	1000	-.228(.071)[.068]	-.258(.084)[.084]{.088}	-.271(.087)[.085]	-.272(.088)[.085]
-.50	100	-.556(.216)[.209]	-.581(.312)[.301]{.299}	-.712(.409)[.350]	-.706(.404)[.347]
	250	-.464(.137)[.132]	-.526(.185)[.183]{.179}	-.576(.202)[.186]	-.579(.204)[.188]
	500	-.439(.113)[.095]	-.514(.129)[.128]{.124}	-.543(.137)[.130]	-.544(.138)[.131]
	1000	-.423(.101)[.066]	-.506(.089)[.089]{.088}	-.520(.092)[.090]	-.521(.092)[.090]

Table 3: Cont'd

λ_0	n	QMLE	MQMLE	RGMM	ORGMM
DGP 2: $\beta_0 = (.3, .1, .1)'$					
.50	100	.361(.206)[.152]	.453(.150)[.143]{.137}	.426(.251)[.240]	.518(.396)[.396]
	250	.435(.103)[.080]	.489(.078)[.077]{.070}	.469(.085)[.079]	.510(.185)[.185]
	500	.453(.070)[.052]	.496(.050)[.050]{.049}	.485(.053)[.051]	.502(.113)[.113]
	1000	.460(.054)[.037]	.497(.036)[.036]{.035}	.492(.038)[.037]	.494(.042)[.042]
.25	100	.098(.241)[.187]	.197(.202)[.194]{.186}	.134(.269)[.242]	.230(.459)[.459]
	250	.176(.131)[.108]	.229(.116)[.114]{.109}	.199(.128)[.117]	.231(.219)[.218]
	500	.200(.086)[.070]	.239(.075)[.074]{.071}	.222(.080)[.075]	.234(.113)[.112]
	1000	.215(.062)[.052]	.246(.055)[.055]{.051}	.238(.057)[.055]	.239(.062)[.061]
.00	100	-.144(.254)[.209]	-.064(.257)[.249]{.241}	-.154(.314)[.273]	-.029(.573)[.573]
	250	-.052(.127)[.116]	-.017(.132)[.131]{.129}	-.054(.146)[.136]	-.015(.267)[.266]
	500	-.032(.091)[.085]	-.014(.098)[.097]{.090}	-.036(.105)[.099]	-.024(.119)[.116]
	1000	-.018(.063)[.060]	-.009(.069)[.069]{.065}	-.020(.072)[.069]	-.014(.082)[.081]
-.25	100	-.354(.235)[.211]	-.311(.283)[.276]{.265}	-.423(.348)[.302]	-.320(.534)[.529]
	250	-.264(.131)[.130]	-.268(.164)[.163]{.159}	-.312(.180)[.168]	-.278(.271)[.269]
	500	-.241(.090)[.089]	-.260(.110)[.109]{.106}	-.286(.117)[.111]	-.269(.136)[.135]
	1000	-.228(.067)[.064]	-.257(.078)[.078]{.077}	-.270(.081)[.078]	-.266(.092)[.091]
-.50	100	-.543(.218)[.214]	-.559(.308)[.302]{.296}	-.696(.424)[.376]	-.621(.616)[.604]
	250	-.468(.138)[.135]	-.532(.186)[.183]{.179}	-.583(.203)[.186]	-.563(.248)[.240]
	500	-.444(.113)[.098]	-.520(.129)[.128]{.122}	-.549(.138)[.129]	-.538(.161)[.156]
	1000	-.420(.104)[.066]	-.503(.086)[.086]{.087}	-.517(.088)[.086]	-.512(.101)[.100]
DGP 3: $\beta_0 = (.3, .1, .1)'$					
.50	100	.378(.186)[.140]	.470(.131)[.127]{.114}	.439(.225)[.217]	.575(.428)[.421]
	250	.434(.100)[.076]	.487(.071)[.070]{.074}	.467(.080)[.073]	.536(.255)[.253]
	500	.450(.074)[.055]	.492(.052)[.051]{.049}	.481(.056)[.052]	.539(.229)[.226]
	1000	.460(.055)[.037]	.497(.035)[.035]{.033}	.491(.036)[.035]	.523(.168)[.167]
.25	100	.109(.217)[.165]	.210(.173)[.168]{.160}	.151(.252)[.232]	.286(.518)[.517]
	250	.183(.120)[.099]	.235(.103)[.102]{.099}	.207(.114)[.106]	.310(.398)[.394]
	500	.205(.081)[.067]	.243(.069)[.069]{.066}	.227(.074)[.070]	.287(.286)[.284]
	1000	.215(.058)[.046]	.246(.048)[.048]{.047}	.237(.050)[.048]	.265(.179)[.179]
.00	100	-.144(.241)[.194]	-.063(.235)[.227]{.199}	-.144(.329)[.296]	.056(.696)[.694]
	250	-.051(.123)[.112]	-.018(.130)[.129]{.119}	-.054(.144)[.133]	.094(.551)[.543]
	500	-.027(.084)[.079]	-.008(.091)[.090]{.089}	-.030(.098)[.093]	.032(.337)[.336]
	1000	-.015(.058)[.056]	-.006(.065)[.064]{.061}	-.017(.067)[.065]	.020(.210)[.209]
-.25	100	-.355(.231)[.205]	-.313(.273)[.265]{.250}	-.432(.357)[.307]	-.193(.780)[.778]
	250	-.267(.129)[.128]	-.272(.162)[.160]{.151}	-.317(.180)[.167]	-.183(.540)[.536]
	500	-.240(.087)[.086]	-.259(.106)[.106]{.100}	-.285(.114)[.108]	-.202(.376)[.373]
	1000	-.224(.068)[.063]	-.254(.075)[.075]{.073}	-.267(.078)[.076]	-.213(.253)[.251]
-.50	100	-.544(.209)[.204]	-.557(.290)[.284]{.279}	-.684(.447)[.407]	-.442(.904)[.903]
	250	-.467(.139)[.135]	-.526(.179)[.177]{.168}	-.577(.196)[.180]	-.464(.523)[.522]
	500	-.433(.119)[.099]	-.506(.123)[.123]{.119}	-.535(.130)[.125]	-.412(.483)[.475]
	1000	-.423(.107)[.074]	-.504(.086)[.086]{.083}	-.519(.088)[.086]	-.466(.257)[.255]

Table 4: Empirical Mean(rmse)[sd]{$\hat{\text{sd}}$} of Estimators of β for SAR Model
Cases of Consistent QMLEs

λ_0	n	β_0	QMLE	MQMLE	RGMM	ORGMM
			DGP 1: Constant Circular Neighbours (REG-1), $\beta_0 = (3, 1, 1)'$			
.5	100	3	3.220(.644)[.606]{.592}	3.166(.708)[.688]{.691}	3.192(.733)[.707]	3.129(.797)[.786]
		1	1.006(.131)[.131]{.123}	0.992(.153)[.152]{.143}	0.989(.152)[.152]	0.988(.152)[.152]
		1	1.003(.201)[.201]{.203}	0.990(.229)[.228]{.222}	0.983(.228)[.228]	0.981(.229)[.229]
	250	3	3.089(.396)[.386]{.392}	3.051(.388)[.385]{.369}	3.069(.395)[.389]	3.040(.437)[.435]
		1	0.999(.096)[.096]{.093}	0.999(.096)[.096]{.093}	0.996(.096)[.096]	0.996(.096)[.096]
		1	1.003(.138)[.138]{.134}	1.004(.149)[.149]{.144}	1.002(.149)[.149]	1.002(.149)[.149]
	500	3	3.039(.264)[.261]{.276}	3.019(.261)[.260]{.253}	3.030(.264)[.263]	3.013(.290)[.290]
		1	1.000(.068)[.068]{.068}	0.996(.070)[.070]{.070}	0.995(.070)[.070]	0.995(.070)[.070]
		1	0.999(.106)[.106]{.104}	0.998(.106)[.106]{.104}	0.997(.106)[.106]	0.997(.106)[.106]
-.5	100	3	3.047(.357)[.353]{.360}	3.011(.356)[.355]{.339}	3.041(.362)[.360]	3.024(.400)[.399]
		1	0.994(.130)[.130]{.123}	0.994(.157)[.157]{.149}	0.988(.157)[.157]	0.988(.158)[.158]
		1	0.995(.226)[.226]{.222}	0.996(.227)[.227]{.222}	0.988(.226)[.226]	0.987(.227)[.227]
	250	3	3.026(.221)[.220]{.230}	3.011(.220)[.220]{.214}	3.024(.221)[.220]	3.016(.246)[.245]
		1	0.999(.098)[.098]{.100}	0.995(.093)[.093]{.094}	0.992(.094)[.093]	0.992(.094)[.093]
		1	1.002(.130)[.130]{.135}	0.992(.143)[.143]{.144}	0.989(.143)[.143]	0.990(.144)[.143]
	500	3	3.001(.157)[.157]{.166}	2.993(.158)[.157]{.152}	3.000(.158)[.158]	2.993(.174)[.174]
		1	0.998(.067)[.067]{.068}	0.998(.067)[.067]{.070}	0.997(.067)[.067]	0.997(.067)[.067]
		1	0.999(.104)[.104]{.103}	0.999(.104)[.104]{.103}	0.997(.104)[.104]	0.998(.104)[.104]
			DGP 2: Constant Circular Neighbours (REG-1), $\beta_0 = (3, 1, 1)'$			
.5	100	3	3.207(.597)[.560]{.570}	3.117(.641)[.631]{.645}	3.150(.706)[.690]	3.071(.843)[.840]
		1	1.007(.154)[.154]{.148}	1.007(.154)[.154]{.148}	1.003(.154)[.154]	1.003(.151)[.151]
		1	1.000(.207)[.207]{.198}	0.999(.220)[.220]{.211}	0.993(.220)[.220]	0.991(.217)[.217]
	250	3	3.078(.380)[.372]{.345}	3.041(.372)[.370]{.345}	3.057(.377)[.372]	3.029(.512)[.512]
		1	1.004(.096)[.096]{.092}	1.004(.096)[.096]{.092}	1.001(.095)[.095]	1.001(.095)[.095]
		1	0.993(.141)[.141]{.132}	1.010(.146)[.146]{.141}	1.007(.146)[.146]	1.007(.145)[.145]
	500	3	3.028(.254)[.253]{.229}	3.009(.252)[.252]{.245}	3.020(.254)[.253]	2.998(.357)[.357]
		1	1.001(.067)[.067]{.068}	0.996(.071)[.070]{.069}	0.995(.071)[.070]	0.995(.070)[.070]
		1	0.999(.100)[.100]{.097}	1.002(.108)[.108]{.103}	1.001(.108)[.108]	1.000(.108)[.108]
-.5	100	3	3.044(.326)[.323]{.310}	3.010(.324)[.324]{.316}	3.039(.331)[.329]	3.021(.450)[.449]
		1	0.997(.154)[.154]{.141}	0.999(.154)[.154]{.140}	0.992(.154)[.153]	0.993(.153)[.153]
		1	0.999(.235)[.235]{.217}	1.000(.235)[.235]{.218}	0.992(.234)[.234]	0.990(.231)[.231]
	250	3	3.012(.205)[.205]{.201}	2.997(.205)[.205]{.206}	3.010(.206)[.206]	3.002(.281)[.281]
		1	1.000(.097)[.097]{.093}	1.001(.097)[.097]{.093}	0.998(.097)[.097]	0.999(.097)[.097]
		1	0.997(.147)[.147]{.141}	0.998(.147)[.147]{.142}	0.994(.147)[.147]	0.995(.146)[.145]
	500	3	3.010(.148)[.148]{.101}	3.002(.148)[.148]{.150}	3.009(.148)[.148]	3.002(.207)[.207]
		1	1.001(.070)[.070]{.067}	0.995(.069)[.069]{.069}	0.994(.069)[.069]	0.994(.069)[.069]
		1	1.000(.104)[.104]{.103}	1.000(.104)[.104]{.103}	0.998(.104)[.104]	0.999(.103)[.103]

Table 4: Cont'd

λ_0	n	β_0	QMLE	MQMLE	RGMM	ORGMM
DGP 1: Queen Contiguity (REG-1), $\beta_0 = (.3, .1, .1)'$						
.5	100	.3	.338(.154)[.149]{.139}	.323(.146)[.145]{.137}	.328(.154)[.151]	.306(.167)[.167]
		.1	.094(.163)[.163]{.159}	.094(.163)[.163]{.169}	.093(.162)[.162]	.092(.165)[.165]
		.1	.100(.204)[.204]{.195}	.100(.204)[.204]{.195}	.099(.202)[.202]	.100(.202)[.202]
	250	.3	.310(.082)[.081]{.082}	.303(.080)[.080]{.079}	.307(.081)[.081]	.300(.081)[.081]
		.1	.109(.096)[.096]{.096}	.109(.096)[.096]{.096}	.109(.096)[.095]	.108(.096)[.096]
		.1	.101(.139)[.139]{.134}	.096(.141)[.141]{.139}	.096(.141)[.141]	.096(.140)[.140]
	500	.3	.308(.060)[.059]{.059}	.304(.059)[.058]{.056}	.306(.064)[.064]	.302(.064)[.064]
		.1	.101(.067)[.067]{.068}	.101(.067)[.067]{.068}	.101(.067)[.067]	.100(.067)[.067]
		.1	.102(.100)[.100]{.098}	.102(.100)[.100]{.098}	.102(.100)[.100]	.101(.100)[.100]
-.5	100	.3	.306(.109)[.109]{.106}	.301(.108)[.108]{.104}	.305(.110)[.109]	.304(.110)[.110]
		.1	.100(.167)[.167]{.157}	.100(.168)[.168]{.159}	.099(.166)[.166]	.097(.168)[.168]
		.1	.087(.195)[.194]{.185}	.084(.199)[.198]{.189}	.082(.196)[.195]	.082(.196)[.195]
	250	.3	.303(.069)[.069]{.069}	.303(.069)[.069]{.068}	.305(.069)[.069]	.306(.069)[.069]
		.1	.097(.099)[.098]{.095}	.107(.100)[.100]{.095}	.106(.100)[.100]	.106(.100)[.099]
		.1	.096(.138)[.138]{.134}	.106(.138)[.138]{.133}	.105(.138)[.138]	.105(.138)[.138]
	500	.3	.301(.048)[.048]{.048}	.297(.049)[.049]{.048}	.298(.049)[.049]	.298(.049)[.049]
		.1	.100(.069)[.069]{.067}	.101(.069)[.069]{.067}	.101(.069)[.069]	.101(.069)[.069]
		.1	.100(.097)[.097]{.098}	.100(.097)[.097]{.098}	.100(.097)[.097]	.100(.097)[.097]
DGP 2: Queen Contiguity (REG-1), $\beta_0 = (.3, .1, .1)'$						
.5	100	.3	.327(.136)[.133]{.128}	.311(.129)[.129]{.120}	.318(.134)[.133]	.251(.234)[.229]
		.1	.103(.161)[.161]{.153}	.103(.161)[.161]{.152}	.103(.161)[.161]	.102(.161)[.161]
		.1	.103(.194)[.194]{.189}	.094(.194)[.194]{.180}	.092(.193)[.193]	.093(.192)[.192]
	250	.3	.311(.080)[.079]{.087}	.304(.078)[.078]{.078}	.308(.079)[.079]	.280(.111)[.110]
		.1	.104(.095)[.095]{.093}	.108(.097)[.097]{.093}	.107(.097)[.097]	.106(.095)[.095]
		.1	.096(.130)[.130]{.132}	.096(.130)[.130]{.132}	.096(.129)[.129]	.096(.129)[.129]
	500	.3	.307(.057)[.057]{.064}	.305(.058)[.058]{.056}	.306(.064)[.063]	.292(.070)[.069]
		.1	.101(.069)[.069]{.067}	.101(.069)[.069]{.067}	.101(.069)[.069]	.100(.068)[.068]
		.1	.104(.102)[.102]{.098}	.094(.101)[.101]{.098}	.094(.101)[.101]	.092(.100)[.099]
-.5	100	.3	.306(.109)[.109]{.110}	.301(.108)[.108]{.103}	.306(.109)[.109]	.304(.111)[.111]
		.1	.104(.171)[.171]{.162}	.104(.172)[.172]{.164}	.103(.170)[.170]	.103(.159)[.159]
		.1	.101(.194)[.194]{.181}	.089(.194)[.194]{.181}	.088(.192)[.191]	.084(.181)[.180]
	250	.3	.300(.069)[.069]{.072}	.302(.067)[.067]{.066}	.304(.067)[.067]	.303(.070)[.070]
		.1	.103(.095)[.095]{.093}	.103(.095)[.095]{.093}	.102(.095)[.094]	.101(.092)[.092]
		.1	.101(.133)[.133]{.132}	.095(.138)[.138]{.130}	.094(.138)[.138]	.093(.133)[.133]
	500	.3	.299(.048)[.048]{.051}	.298(.048)[.048]{.048}	.299(.048)[.048]	.299(.049)[.049]
		.1	.102(.067)[.067]{.068}	.102(.067)[.067]{.068}	.101(.067)[.067]	.100(.066)[.066]
		.1	.099(.099)[.099]{.096}	.103(.103)[.103]{.098}	.103(.102)[.102]	.103(.101)[.101]

Table 5: Empirical Mean(rmse)[sd]{$\hat{\text{sd}}$} of Estimators of β for SAR Model
Case I of Inconsistent QMLEs: Circular Neighbours (REG-1)

λ_0	n	β_0	QMLE	MQMLE	RGMM	ORGMM
			DGP 1: $\beta_0 = (3,1,1)'$			
.5	100	3	3.398(.598)[.719]	3.116(.596)[.607]{.594}	3.145(.641)[.624]	3.104(.679)[.671]
		1	1.001(.125)[.125]	0.997(.125)[.125]{.118}	0.993(.125)[.125]	0.993(.126)[.126]
		1	0.999(.190)[.190]	0.992(.189)[.189]{.188}	0.986(.188)[.187]	0.987(.187)[.187]
	250	3	3.254(.346)[.429]	3.055(.350)[.355]{.349}	3.067(.351)[.345]	3.048(.370)[.367]
		1	1.001(.076)[.076]	0.998(.076)[.076]{.073}	0.997(.076)[.076]	0.997(.076)[.076]
		1	1.011(.125)[.125]	1.004(.124)[.124]{.119}	1.002(.124)[.124]	1.002(.124)[.124]
	500	3	3.219(.263)[.342]	3.024(.265)[.266]{.262}	3.030(.266)[.264]	3.021(.281)[.280]
		1	1.006(.054)[.055]	1.000(.054)[.054]{.056}	0.999(.054)[.054]	0.999(.055)[.055]
		1	1.008(.090)[.090]	1.002(.089)[.089]{.089}	1.001(.089)[.089]	1.001(.089)[.089]
-.5	100	3	2.897(.206)[.231]	2.986(.259)[.259]{.270}	2.981(.232)[.231]	2.993(.245)[.245]
		1	1.003(.127)[.127]	0.999(.127)[.127]{.120}	0.996(.127)[.127]	0.995(.127)[.127]
		1	1.014(.191)[.191]	1.003(.192)[.192]{.194}	0.996(.192)[.192]	0.993(.192)[.192]
	250	3	2.898(.134)[.169]	3.010(.177)[.177]{.166}	3.000(.146)[.146]	3.003(.154)[.154]
		1	1.005(.072)[.073]	0.996(.072)[.072]{.074}	0.995(.073)[.073]	0.995(.073)[.073]
		1	1.001(.122)[.122]	0.996(.121)[.121]{.119}	0.995(.121)[.121]	0.995(.121)[.121]
	500	3	2.887(.101)[.152]	3.011(.136)[.137]{.135}	3.009(.115)[.115]	3.011(.121)[.120]
		1	1.003(.055)[.055]	1.000(.055)[.055]{.055}	0.999(.055)[.055]	0.999(.055)[.055]
		1	1.002(.089)[.089]	0.995(.089)[.089]{.088}	0.994(.089)[.089]	0.993(.089)[.089]
			DGP 2: $\beta_0 = (3,1,1)'$			
.5	100	3	3.374(.572)[.683]	3.104(.568)[.578]{.563}	3.136(.623)[.607]	3.092(.767)[.762]
		1	1.009(.122)[.122]	1.005(.122)[.122]{.115}	1.001(.122)[.122]	1.001(.121)[.121]
		1	0.995(.193)[.193]	0.988(.192)[.193]{.182}	0.982(.192)[.191]	0.983(.192)[.191]
	250	3	3.229(.325)[.397]	3.030(.327)[.328]{.330}	3.045(.321)[.318]	3.021(.399)[.399]
		1	1.000(.073)[.073]	0.997(.073)[.073]{.072}	0.995(.074)[.073]	0.995(.074)[.074]
		1	1.013(.118)[.118]	1.006(.117)[.118]{.117}	1.004(.118)[.118]	1.005(.118)[.118]
	500	3	3.200(.261)[.329]	3.003(.262)[.262]{.259}	3.013(.265)[.264]	3.005(.343)[.343]
		1	1.007(.054)[.055]	1.001(.054)[.054]{.055}	1.000(.054)[.054]	1.000(.054)[.054]
		1	1.006(.089)[.089]	1.001(.088)[.088]{.087}	1.000(.088)[.088]	1.000(.088)[.088]
-.5	100	3	2.907(.209)[.229]	3.002(.260)[.260]{.273}	2.992(.239)[.239]	2.994(.265)[.265]
		1	0.997(.125)[.125]	0.993(.124)[.124]{.119}	0.990(.125)[.124]	0.991(.124)[.124]
		1	1.016(.198)[.199]	1.003(.199)[.199]{.195}	0.997(.200)[.200]	0.998(.199)[.199]
	250	3	2.892(.135)[.173]	3.000(.168)[.168]{.161}	2.995(.145)[.145]	2.996(.169)[.168]
		1	1.010(.075)[.076]	1.001(.075)[.075]{.072}	1.000(.076)[.076]	1.000(.076)[.076]
		1	0.996(.122)[.122]	0.991(.121)[.121]{.116}	0.989(.121)[.121]	0.990(.121)[.121]
	500	3	2.875(.101)[.161]	2.997(.133)[.133]{.129}	2.994(.113)[.113]	2.991(.137)[.137]
		1	1.007(.056)[.057]	1.004(.056)[.056]{.055}	1.003(.056)[.056]	1.003(.056)[.056]
		1	1.010(.090)[.090]	1.002(.090)[.090]{.088}	1.001(.090)[.090]	1.001(.090)[.090]

Table 6: Empirical Mean(rmse)[sd]{$\hat{\text{sd}}$} of Estimators of β for SAR Model
Case II of Inconsistent QMLEs: Group Interaction (REG-2)

λ_0	n	β_0	QMLE	MQMLE	RGMM	ORGMM
			DGP 1: $\beta_0 = (3, 1, 1)'$			
.5	100	3	3.493(.795)[.623]	3.146(.645)[.628]{.599}	3.207(.698)[.667]	3.196(.714)[.687]
		1	1.131(.253)[.217]	1.036(.221)[.218]{.205}	1.043(.237)[.233]	1.043(.239)[.235]
		1	1.096(.272)[.254]	1.015(.245)[.244]{.247}	1.019(.260)[.260]	1.019(.262)[.261]
	250	3	3.239(.423)[.348]	3.041(.358)[.355]{.349}	3.074(.375)[.367]	3.054(.397)[.394]
		1	1.059(.160)[.149]	1.008(.149)[.149]{.142}	1.012(.151)[.151]	1.008(.155)[.155]
		1	1.058(.160)[.149]	1.007(.149)[.149]{.139}	1.011(.152)[.151]	1.008(.155)[.155]
	500	3	3.173(.291)[.234]	3.017(.237)[.236]{.239}	3.038(.245)[.242]	3.027(.258)[.256]
		1	1.045(.101)[.090]	1.002(.090)[.090]{.091}	1.006(.091)[.091]	1.003(.093)[.093]
		1	1.045(.106)[.096]	1.004(.096)[.096]{.099}	1.008(.097)[.096]	1.005(.099)[.099]
-.5	100	3	3.070(.388)[.382]	3.075(.489)[.483]{.480}	3.104(.493)[.482]	3.097(.521)[.512]
		1	1.011(.168)[.168]	1.011(.183)[.182]{.202}	1.009(.190)[.190]	1.009(.194)[.194]
		1	1.019(.230)[.229]	1.020(.247)[.246]{.245}	1.016(.243)[.242]	1.015(.245)[.245]
	250	3	2.938(.251)[.243]	3.015(.308)[.307]{.301}	3.033(.296)[.294]	3.025(.312)[.310]
		1	0.980(.129)[.127]	0.997(.135)[.135]{.134}	0.998(.136)[.136]	0.997(.139)[.139]
		1	0.982(.127)[.125]	1.000(.134)[.134]{.131}	1.001(.134)[.134]	1.001(.136)[.136]
	500	3	2.918(.189)[.170]	3.013(.216)[.215]{.204}	3.023(.202)[.200]	3.017(.212)[.212]
		1	0.976(.082)[.078]	1.001(.087)[.087]{.083}	1.002(.083)[.083]	1.001(.085)[.085]
		1	0.976(.086)[.083]	1.000(.088)[.088]{.092}	1.001(.087)[.087]	0.999(.089)[.089]
			DGP 2: $\beta_0 = (3, 1, 1)'$			
.5	100	3	3.397(.746)[.631]	3.057(.622)[.620]{.654}	3.088(.693)[.688]	3.027(.786)[.786]
		1	1.106(.239)[.214]	1.012(.213)[.213]{.198}	1.009(.234)[.234]	0.998(.255)[.255]
		1	1.084(.277)[.264]	1.003(.252)[.252]{.239}	0.999(.275)[.275]	0.989(.285)[.285]
	250	3	3.211(.408)[.349]	3.006(.349)[.349]{.333}	3.036(.366)[.364]	2.979(.450)[.449]
		1	1.045(.152)[.146]	0.993(.145)[.144]{.141}	0.996(.148)[.148]	0.984(.165)[.165]
		1	1.046(.153)[.145]	0.993(.144)[.144]{.138}	0.997(.148)[.148]	0.984(.163)[.162]
	500	3	3.172(.287)[.229]	3.016(.230)[.230]{.235}	3.036(.238)[.235]	3.005(.303)[.303]
		1	1.049(.102)[.090]	1.005(.090)[.090]{.091}	1.009(.091)[.091]	1.001(.105)[.105]
		1	1.046(.110)[.100]	1.005(.101)[.101]{.099}	1.008(.101)[.101]	1.001(.112)[.112]
-.5	100	3	3.055(.397)[.394]	3.073(.520)[.515]{.508}	3.096(.508)[.499]	3.031(.598)[.597]
		1	1.016(.174)[.173]	1.020(.197)[.196]{.218}	1.019(.197)[.196]	1.004(.214)[.214]
		1	1.004(.225)[.225]	1.009(.246)[.246]{.260}	1.001(.241)[.241]	0.991(.248)[.248]
	250	3	2.939(.247)[.239]	3.018(.301)[.300]{.392}	3.031(.286)[.284]	2.992(.357)[.357]
		1	0.986(.128)[.127]	1.006(.136)[.136]{.133}	1.005(.137)[.137]	0.997(.149)[.148]
		1	0.986(.123)[.122]	1.005(.132)[.131]{.130}	1.006(.130)[.130]	0.997(.140)[.140]
	500	3	2.912(.195)[.174]	3.003(.216)[.216]{.200}	3.015(.206)[.205]	2.993(.253)[.253]
		1	0.976(.081)[.078]	1.000(.085)[.085]{.083}	1.002(.083)[.083]	0.996(.091)[.091]
		1	0.982(.090)[.088]	1.005(.093)[.093]{.092}	1.007(.094)[.093]	1.002(.100)[.100]

Table 7: Empirical Mean(rmse)[sd] of Estimators of λ and ρ for SARAR(1,1) Model
Case I of Inconsistent QMLEs: Circular Neighbours (REG-1)

Par	QMLE-λ	MQMLE-λ	KP-λ	QMLE-ρ	MQMLE-ρ	KP-ρ
			DGP 1: $\beta_0 = (3, 1, 1)'$			
1-1	.470(.141)[.138]	.472(.197)[.195]	.578(.219)[.204]	.409(.195)[.172]	.446(.237)[.231]	.335(.341)[.299]
	.484(.080)[.078]	.482(.118)[.117]	.528(.109)[.105]	.445(.116)[.102]	.488(.140)[.139]	.479(.180)[.179]
	.487(.065)[.064]	.489(.097)[.097]	.515(.093)[.092]	.454(.088)[.075]	.491(.110)[.109]	.512(.156)[.156]
	.490(.043)[.042]	.495(.060)[.059]	.505(.057)[.057]	.458(.066)[.051]	.497(.070)[.070]	.533(.103)[.097]
1-2	.372(.173)[.116]	.418(.233)[.218]	.494(.143)[.143]	-.307(.249)[.158]	-.505(.252)[.239]	-.507(.244)[.244]
	.411(.109)[.063]	.488(.095)[.094]	.501(.072)[.072]	-.324(.202)[.100]	-.502(.153)[.153]	-.492(.150)[.150]
	.400(.112)[.050]	.498(.071)[.071]	.498(.060)[.060]	-.305(.208)[.072]	-.504(.126)[.125]	-.476(.121)[.119]
	.421(.084)[.030]	.502(.047)[.047]	.499(.035)[.035]	-.321(.186)[.051]	-.506(.109)[.108]	-.470(.083)[.078]
2-1	.280(.144)[.141]	.250(.200)[.200]	.333(.239)[.224]	.374(.208)[.165]	.441(.225)[.217]	.358(.307)[.272]
	.292(.095)[.086]	.253(.128)[.127]	.297(.133)[.124]	.399(.140)[.097]	.470(.135)[.131]	.464(.176)[.172]
	.293(.080)[.067]	.252(.106)[.106]	.276(.105)[.101]	.408(.119)[.075]	.491(.109)[.107]	.499(.146)[.146]
	.287(.057)[.043]	.250(.064)[.064]	.259(.064)[.064]	.421(.093)[.049]	.494(.065)[.065]	.524(.092)[.089]
2-2	.113(.189)[.130]	.233(.188)[.163]	.235(.186)[.186]	-.330(.231)[.156]	-.582(.269)[.249]	-.507(.259)[.259]
	.156(.120)[.074]	.239(.131)[.131]	.248(.092)[.092]	-.337(.188)[.095]	-.503(.209)[.209]	-.484(.151)[.150]
	.140(.125)[.059]	.248(.099)[.099]	.247(.079)[.079]	-.319(.193)[.069]	-.510(.115)[.114]	-.484(.117)[.116]
	.164(.093)[.036]	.250(.052)[.052]	.250(.045)[.045]	-.332(.175)[.047]	-.501(.102)[.101]	-.475(.080)[.076]
3-1	.082(.168)[.147]	.015(.210)[.209]	.080(.236)[.222]	.335(.239)[.172]	.428(.241)[.230]	.367(.292)[.260]
	.090(.124)[.086]	.012(.126)[.125]	.047(.131)[.123]	.373(.160)[.097]	.472(.127)[.124]	.467(.165)[.161]
	.094(.116)[.067]	.006(.099)[.099]	.025(.103)[.100]	.380(.140)[.072]	.495(.093)[.091]	.502(.131)[.131]
	.082(.093)[.043]	.001(.062)[.062]	.009(.064)[.063]	.397(.114)[.048]	.496(.059)[.059]	.526(.088)[.083]
3-2	-.104(.163)[.125]	-.027(.171)[.148]	-.027(.196)[.194]	-.353(.208)[.147]	-.488(.208)[.187]	-.485(.250)[.250]
	-.078(.109)[.076]	-.023(.152)[.150]	-.006(.108)[.108]	-.356(.170)[.091]	-.489(.117)[.116]	-.481(.148)[.147]
	-.086(.106)[.062]	.000(.123)[.123]	-.001(.096)[.096]	-.343(.170)[.065]	-.501(.102)[.102]	-.478(.120)[.117]
	-.071(.081)[.040]	.001(.060)[.059]	-.001(.055)[.055]	-.350(.156)[.045]	-.502(.106)[.104]	-.473(.081)[.077]
4-1	-.126(.183)[.135]	-.219(.194)[.192]	-.189(.210)[.201]	.323(.246)[.170]	.430(.239)[.228]	.395(.258)[.235]
	-.132(.144)[.082]	-.240(.106)[.105]	-.224(.117)[.114]	.363(.169)[.099]	.478(.112)[.109]	.485(.150)[.149]
	-.119(.148)[.068]	-.247(.090)[.090]	-.236(.095)[.093]	.365(.155)[.075]	.490(.085)[.085]	.510(.118)[.118]
	-.131(.127)[.043]	-.247(.057)[.057]	-.242(.055)[.055]	.376(.134)[.049]	.492(.056)[.055]	.520(.079)[.076]
4-2	-.303(.130)[.119]	-.300(.217)[.215]	-.279(.208)[.206]	-.395(.168)[.131]	-.484(.224)[.228]	-.488(.221)[.221]
	-.288(.084)[.075]	-.272(.155)[.154]	-.260(.122)[.121]	-.384(.145)[.086]	-.487(.200)[.199]	-.475(.152)[.150]
	-.289(.069)[.057]	-.249(.106)[.106]	-.255(.098)[.098]	-.378(.137)[.061]	-.508(.101)[.100]	-.472(.115)[.112]
	-.284(.050)[.037]	-.244(.056)[.056]	-.253(.058)[.058]	-.381(.126)[.043]	-.506(.104)[.103]	-.471(.081)[.076]
5-1	-.357(.192)[.128]	-.458(.165)[.160]	-.449(.169)[.161]	.320(.244)[.164]	.438(.201)[.191]	.413(.215)[.197]
	-.373(.146)[.071]	-.491(.082)[.082]	-.481(.088)[.086]	.362(.169)[.097]	.483(.101)[.100]	.488(.126)[.126]
	-.352(.159)[.057]	-.496(.068)[.068]	-.493(.073)[.073]	.357(.160)[.072]	.491(.074)[.074]	.509(.097)[.097]
	-.374(.131)[.037]	-.499(.041)[.041]	-.498(.045)[.045]	.377(.132)[.047]	.497(.047)[.047]	.526(.069)[.064]
5-2	-.478(.104)[.101]	-.523(.180)[.179]	-.518(.189)[.188]	-.437(.140)[.125]	-.490(.215)[.214]	-.491(.217)[.217]
	-.480(.069)[.066]	-.513(.128)[.126]	-.511(.111)[.111]	-.424(.113)[.084]	-.491(.146)[.145]	-.474(.142)[.140]
	-.472(.057)[.050]	-.498(.109)[.109]	-.501(.093)[.093]	-.429(.092)[.059]	-.507(.107)[.106]	-.474(.113)[.110]
	-.478(.040)[.033]	-.499(.054)[.053]	-.502(.056)[.056]	-.424(.086)[.042]	-.500(.077)[.076]	-.470(.078)[.073]

Note: (i) The DGP used: $Y_n = \lambda W_n Y_n + \iota_n \beta_0 + X_{1n}\beta_1 + X_{2n}\beta_2 + \epsilon_n, \quad \epsilon_n = \rho W_n \epsilon_n + v_n$.

(ii) Par = i-j, where '$i = 1, 2, 3, 4, 5$' represents '$\lambda = .5, .25, 0, -.25, -.5$'; '$j = 1, 2$' represents '$\rho = .5, -.5$'. Under each Par setting, $n = 100, 250, 500, 1000$, corresponding to the four rows.

(iii) KP denotes Kelejian and Prucha's (2010) three-step estimator.

Table 7: Cont'd

Par	λ_{QML}	λ_{MQML}	λ_{GS}	ρ_{QML}	ρ_{MQML}	ρ_{GS}
DGP 2: $\beta_0 = (3, 1, 1)'$						
1-1	.471(.138)[.135]	.473(.188)[.186]	.576(.216)[.202]	.404(.195)[.170]	.443(.235)[.228]	.329(.340)[.294]
	.487(.085)[.084]	.489(.123)[.122]	.529(.114)[.110]	.440(.119)[.103]	.477(.144)[.142]	.466(.180)[.177]
	.490(.064)[.063]	.494(.096)[.096]	.521(.092)[.089]	.454(.090)[.077]	.489(.113)[.113]	.513(.158)[.157]
	.493(.042)[.041]	.499(.059)[.059]	.508(.056)[.056]	.455(.068)[.051]	.493(.070)[.070]	.526(.099)[.095]
1-2	.380(.173)[.124]	.474(.153)[.138]	.495(.141)[.141]	-.318(.253)[.175]	-.498(.266)[.252]	-.509(.246)[.246]
	.408(.114)[.067]	.483(.101)[.100]	.495(.072)[.072]	-.320(.212)[.113]	-.492(.200)[.199]	-.483(.156)[.155]
	.399(.116)[.056]	.497(.063)[.063]	.497(.060)[.060]	-.303(.215)[.085]	-.504(.124)[.123]	-.476(.125)[.123]
	.422(.085)[.035]	.502(.038)[.038]	.500(.035)[.035]	-.324(.185)[.056]	-.506(.100)[.100]	-.473(.079)[.074]
2-1	.285(.143)[.139]	.262(.194)[.194]	.349(.241)[.220]	.370(.214)[.170]	.432(.232)[.222]	.345(.317)[.276]
	.280(.092)[.087]	.244(.126)[.126]	.282(.127)[.123]	.407(.136)[.099]	.478(.134)[.132]	.478(.177)[.176]
	.292(.081)[.069]	.253(.105)[.105]	.275(.106)[.103]	.411(.116)[.074]	.483(.103)[.102]	.502(.145)[.145]
	.285(.056)[.043]	.247(.064)[.064]	.257(.063)[.062]	.424(.091)[.049]	.497(.065)[.064]	.529(.096)[.092]
2-2	.120(.186)[.134]	.235(.198)[.175]	.227(.184)[.182]	-.337(.235)[.169]	-.490(.271)[.254]	-.504(.263)[.263]
	.157(.121)[.078]	.247(.109)[.108]	.249(.094)[.094]	-.336(.194)[.104]	-.492(.125)[.124]	-.485(.153)[.152]
	.141(.125)[.062]	.251(.101)[.101]	.247(.081)[.081]	-.318(.197)[.076]	-.502(.129)[.127]	-.479(.121)[.119]
	.163(.096)[.039]	.252(.043)[.043]	.248(.047)[.046]	-.334(.174)[.053]	-.507(.101)[.109]	-.474(.083)[.079]
3-1	.086(.171)[.148]	.033(.210)[.207]	.085(.234)[.218]	.339(.231)[.166]	.419(.235)[.221]	.371(.282)[.250]
	.082(.121)[.089]	.007(.124)[.124]	.037(.130)[.125]	.376(.161)[.103]	.472(.128)[.125]	.476(.166)[.164]
	.092(.114)[.068]	.005(.096)[.096]	.022(.102)[.100]	.382(.138)[.071]	.490(.088)[.087]	.502(.124)[.124]
	.081(.092)[.044]	.001(.061)[.061]	.009(.061)[.060]	.395(.116)[.050]	.493(.058)[.058]	.525(.088)[.085]
3-2	-.087(.156)[.129]	-.026(.205)[.201]	-.012(.199)[.199]	-.367(.196)[.144]	-.479(.211)[.197]	-.492(.238)[.238]
	-.078(.109)[.077]	-.021(.138)[.137]	-.007(.109)[.109]	-.356(.173)[.096]	-.487(.195)[.195]	-.479(.152)[.150]
	-.088(.108)[.062]	-.005(.106)[.106]	-.009(.092)[.091]	-.344(.171)[.071]	-.508(.115)[.115]	-.474(.120)[.117]
	-.070(.080)[.039]	.008(.057)[.057]	.000(.054)[.054]	-.352(.156)[.049]	-.508(.107)[.105]	-.472(.080)[.075]
4-1	-.132(.186)[.144]	-.214(.201)[.198]	-.185(.215)[.205]	.329(.238)[.165]	.428(.227)[.215]	.393(.245)[.221]
	-.132(.148)[.090]	-.237(.119)[.118]	-.219(.120)[.116]	.366(.170)[.105]	.477(.128)[.126]	.481(.150)[.149]
	-.119(.149)[.072]	-.246(.088)[.088]	-.236(.096)[.095]	.368(.153)[.076]	.491(.082)[.082]	.508(.114)[.114]
	-.133(.125)[.045]	-.248(.055)[.055]	-.244(.057)[.056]	.382(.129)[.051]	.497(.054)[.054]	.527(.079)[.074]
4-2	-.297(.132)[.123]	-.287(.206)[.205]	-.280(.207)[.205]	-.401(.175)[.144]	-.462(.232)[.235]	-.492(.233)[.233]
	-.288(.083)[.074]	-.278(.107)[.104]	-.261(.115)[.115]	-.387(.146)[.092]	-.479(.157)[.156]	-.480(.151)[.150]
	-.289(.071)[.060]	-.253(.100)[.100]	-.253(.102)[.102]	-.381(.135)[.064]	-.508(.117)[.117]	-.480(.118)[.116]
	-.283(.050)[.038]	-.252(.053)[.053]	-.254(.058)[.058]	-.381(.127)[.045]	-.509(.107)[.106]	-.470(.082)[.076]
5-1	-.373(.186)[.135]	-.472(.154)[.151]	-.459(.163)[.158]	.327(.249)[.180]	.445(.205)[.198]	.414(.223)[.206]
	-.380(.144)[.080]	-.492(.082)[.082]	-.485(.090)[.089]	.369(.167)[.103]	.484(.104)[.102]	.492(.127)[.127]
	-.355(.158)[.063]	-.498(.066)[.066]	-.491(.068)[.067]	.362(.156)[.074]	.495(.070)[.070]	.510(.093)[.092]
	-.375(.132)[.042]	-.498(.043)[.043]	-.498(.044)[.044]	.377(.133)[.051]	.495(.048)[.048]	.524(.068)[.063]
5-2	-.473(.120)[.117]	-.520(.191)[.190]	-.518(.198)[.198]	-.436(.156)[.143]	-.483(.226)[.226]	-.483(.229)[.229]
	-.476(.072)[.068]	-.505(.109)[.108]	-.507(.112)[.112]	-.427(.113)[.086]	-.485(.146)[.146]	-.477(.146)[.144]
	-.472(.059)[.052]	-.502(.107)[.107]	-.500(.096)[.096]	-.431(.093)[.062]	-.503(.107)[.107]	-.481(.109)[.107]
	-.480(.040)[.034]	-.498(.063)[.063]	-.507(.057)[.057]	-.422(.089)[.042]	-.505(.075)[.074]	-.469(.081)[.074]