10-2010

Bias-Corrected Estimation for Spatial Autocorrelation

Zhenlin YANG
Singapore Management University, zlyang@smu.edu.sg

SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

Bias-Corrected Estimation for Spatial Autocorrelation

Zhenlin Yang

October 2010

Paper No. 12-2010

Bias-Corrected Estimation for Spatial Autocorrelation[1]

Zhenlin Yang

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

emails: zlyang@smu.edu.sg

October, 2010

Abstract

The biasedness issue arising from the maximum likelihood estimation of the spatial autoregressive model (SAR) is further investigated under a broader set-up than that in Bao and Ullah (2007a). A major difficulty in analytically evaluating the expectations of ratios of quadratic forms is overcome by a simple bootstrap procedure. With that, the corrections on bias and variance of the spatial estimator can easily be made up to third-order, and once this is done, the estimators of other model parameters become nearly unbiased. Compared with the analytical approach, the new approach is much simpler, and can easily be extended to other models of a similar structure. Extensive Monte Carlo results show that the new approach performs excellently in general.

Key Words: Third-order bias; Third-order variance; Bootstrap; Concentrated estimating equation; Monte Carlo; Quasi-MLE; Spatial layout.

JEL Classification: C10, C21

1 Introduction

The mixed regressive, spatial autoregressive (SAR) model takes the following form

$$Y_n = \lambda W_n Y_n + X_n \beta + u_n \tag{1}$$

where n is the total number of spatial units, Y_n is an $n \times 1$ vector of observations on these spatial units, X_n is an $n \times k$ matrix whose rows are values of constant regressors, W_n is a specified $n \times n$ spatial weights matrix, and u_n is an n-dimensional vector of independent and identically distributed (iid) disturbances of zero mean and finite variance σ^2. λ is the

[1]I benefited from the IVth World Conference of the Spatial Econometrics Association, June 9-12, 2010, Chicago, and the seminar at the Singapore Management University, October, 2009. I am grateful to the support from a research grant (Grant number: C244/MSS9E005) from Singapore Management University and the research assistance from Yan Shen.

scalar spatial parameter, and β is a $p \times 1$ vector of regression coefficient. When there are no regressors (X_n) in the model, the SAR model becomes a pure SAR process.[2]

One popular method for estimating the SAR model is the maximum likelihood (ML) or quasi-maximum likelihood (QML) (Ord, 1975; Smirnov and Anselin, 2001; Lee, 2004a). Let $\theta = (\lambda, \beta', \sigma^2)$ with θ_0 being its true value. Let $A_n(\lambda) = I_n - \lambda W_n$ with I_n being an $n \times n$ identity matrix. If the errors are exactly normal, we have the true log-likelihood function,

$$\ell_n(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \log|A_n(\lambda)| - \frac{1}{2\sigma^2}\left[A_n(\lambda)Y_n - X_n\beta\right]'\left[A_n(\lambda)Y_n - X_n\beta\right]. \quad (2)$$

Maximizing $\ell(\theta)$ gives the ML estimator (MLE) of θ. If the errors are not exactly normal as assumed in this paper, $\ell(\theta)$ can still be used as a working log-likelihood called quasi-log-likelihood and maximizing it would still produce a consistent estimator of θ provided certain regularity conditions are satisfied (Lee, 2004a). The resulted estimator is called the quasi-maximum likelihood estimator (QMLE). Now, given λ, $\ell(\theta)$ can be partially maximized, which gives the constrained QMLEs of β and σ^2, respectively,

$$\hat{\beta}_n(\lambda) = (X_n'X_n)^{-1}X_n'A_n(\lambda)Y_n, \quad (3)$$

$$\hat{\sigma}_n^2(\lambda) = \frac{1}{n}Y_n'A_n'(\lambda)M_nA_n(\lambda)Y_n, \quad (4)$$

where $M_n = I_n - X_n(X_n'X_n)^{-1}X_n'$. These lead to the concentrated log-likelihood of λ as

$$\ell_n^c(\lambda) = -\frac{n}{2}[\log(2\pi) + 1] - \frac{n}{2}\log\hat{\sigma}_n^2(\lambda) + \log|A_n(\lambda)| \quad (5)$$

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

Lee (2004a) gives a detailed study on the asymptotic properties of QMLE $\hat{\theta}_n$. In particular, he showed that the QMLEs of β and λ are $\sqrt{n}$-consistent if each spatial unit depends on fixed number of neighbors, otherwise they are $\sqrt{n/h_n}$-consistent if the number of neighbors is of order h_n such that as $n \to \infty$, $h_n \to \infty$ and $h_n/n \to 0$. The QMLE of σ^2 is always $\sqrt{n}$-consistent. Lee's work sets up a solid theoretical foundation for us to study further the higher-order properties of the QMLE of the SAR Model. Although the SAR model can also

[2]Ever since its introduction by Cliff and Ord (1973, 1981), the SAR model has become very popular in modeling the cross-sectional dependence induced by interactions among economic agents, such as neighborhood effects, copy-catting, spillover effects, and peer-group effects. See, among the others, Anselin (1988, 2001), Case (1991), Case, et al. (1993), Cressie (1993), Besley and Case (1995), Brueckner (1998), Anselin and Bera (1998), Bao and Ullah (2007a), Bell and Bockstael (2000), Bertrand, et al. (2000), Topa (2001), Lee (2003, 2004a, 2007a,b), Mynbaev and Ullah (2008), and Robinson (2010), for an account on various SAR-related theoretical issues and economics applications.

be estimated by other methods such as the two-stage least squares (2SLS), and the method of moments (MM) or GMM (Kelejian and Prucha, 1998, 1999; Lee, 2003, 2007a), that the QML estimation is applicable to the pure SAR process but the GMM is not, and that the MLE is more efficient than MM or GMM estimators (Lee, 2004a, 2007a) provide strong reasons for further studies on the QML estimation of the SAR model.[3]

While the QMLE of the SAR model enjoys the good properties, it has been recognized that the it can be quite biased (see, e.g., Lee (2004a, 2007a) and Bao and Ullah (2007a)), and there are no general methods available for correcting the bias, except Bao and Ullah (2007a) who provided analytical formulas for the second-order bias and mean squared error (MSE) of the QMLE of λ for the pure SAR model under the assumption that u_n is normally distributed. In dealing with the case where σ^2 is unknown, Bao and Ullah (2007a, p.400) advocate the use of concentrated likelihood function of the spatial parameter as (i) it simplifies the maximization procedure substantially, and (ii) it also simplifies the derivations for the higher-order results since it is much easier to work with a scalar case than a vector. We take this path and stress further that these simplifications are even greater if the SAR model involves exogenous regressors, in particular in deriving the higher-order approximations to bias and variance. However, with the general model specified in (1) containing regressors and/or with nonnormal errors, the analytical approach of Bao and Ullah (2007a) in finding the expectations of ratios of various quadratic forms runs into difficulty. Recently, Bao (2010) made an attempt to study this problem jointly using the full likelihood function of (1) and has obtained some interesting results on second-order bias of $\hat{\theta}_n$, but apparently this approach runs into difficulties in MSE approximations due to its high complexity. See Ullah (2004) for a general account on the finite sample econometrics.

In this paper, we follow Bao and Ullah (2007a) and propose to tackle the biasedness issue for a general SAR model using the concentrated quasi-likelihood function. To overcome the difficulty in evaluating the expectations of various ratios of quadratic forms in the expansions, we introduce a simple bootstrap procedure which does not require the repeated re-estimations of the model parameters. Some detailed arguments might be helpful for motivating our approach to bias-correction.

[3] An issue related to the QML estimation is on computation. Maximization of (5) looks simple as it is only one dimensional. However, it involves the computation of the determinant $|A_n(\lambda)|$ at each possible value of λ, which can be a burden when n is large. This computational burden can be alleviated by using the following results, (i) if W_n is symmetric or symmetric before row-normalization, then $\log|A_n(\lambda)| = \sum_{i=1}^{n} \log(1 - \lambda\omega_i)$, where ω_i's are the eigenvalues of W_n (Ord, 1975), and (ii) if W_n is intrinsically asymmetric, the result of (i) is applicable through $\log|A_n(\lambda)| = \frac{1}{2}\log|A_n(\lambda)A_n(\lambda)'|$ (Pace and Barry, 1997). When n is very large, the method of Smirnov and Anselin (2001) can be followed.

First, note that if λ in the SAR model is known, then the model becomes essentially a regular linear regression model. Thus, the estimation of β coefficients given λ would be unbiased, and so is the estimation of the error variance σ^2 if mean-squared-error is used for estimating σ^2 instead of QMLE. This means that the biasedness problem for the SAR model centers at the estimation of the spatial parameter λ, and the bias and variance corrections may only be necessary for the QMLE of λ.[4] A multidimensional problem is thus reduced to a scalar one, which greatly simplifies the higher-order stochastic expansions. Yet still, for these expansions to be of a general practical value, they must be supplemented with simple ways for evaluating the expectations of ratios of quadratic forms. Noting that these ratios are all functions of the parameter vector θ and the error vector u_n with iid elements, naturally, their expectations can be bootstrapped (see Efron, 1979).

To summarize, the proposed approach is hybrid – combining stochastic expansions based on concentrated quasi-score function and bootstrap, with the former providing tractable third-order approximations to the bias and variance of the QMLE $\hat{\lambda}_n$, and the latter making these expansions practically implementable. The proposed approach is shown to be very effective in removing bias and improving inferences, and once the QMLE $\hat{\lambda}_n$ is bias-corrected, the QMLEs of other model parameters almost automatically become unbiased. Compared with the analytical approach of Bao and Ullah (2007a) and Bao (2010), the proposed approach is much simpler, and can easily be extended to other models of a similar structure. For example, in a linear regression model with a response transformation (Box and Cox, 1964), the transformation parameter is the source of bias in the estimation, but given that, the estimation of other parameters does not incur bias. Our approach is applicable to this model for bias correction, but the approach of Bao (2010) is not as the expectations of the quantities appeared in the expansions are not available. Other good examples of this type of models may be the dynamic regressions or dynamic panel regressions.

To assess the performance of the proposed method in correcting the bias of the QMLE $\hat{\lambda}_n$ and the impact of this bias correction on the estimation of other model parameters as well as the usual inferences, a series of Monte Carlo experiments are conducted. The results show that the new approach performs excellently in general, with the bias-corrected QMLE of λ clearly outperforming the original QMLE, and a bias-adjusted QMLE using the formulas of Bao and Ullah (2007a). With the use of the bias-corrected estimator of λ, the corresponding estimators of β and σ^2 become nearly unbiased, and the performance of

[4]Lee (2007a) made a similar remark based on his Monte Carlo results. The same arguments may hold for other models of similar feature such linear regressions with response transformation, dynamic regression, dynamic panel data regression, etc.

the usual regression t-ratios improved. Monte Carlo results further reveal that the original QMLE and the QMLE bias-corrected using the formulas of Bao and Ullah (2007a) depends heavily on the following factors: (a) the spatial layout, in particular whether the number of neighbors grows with n, (b) the error standard deviation σ, and (c) the value of λ. Some anomalies are found in the Monte Carlo results of Bao and Ullah (2007a) and Lee (2004a). Their Monte Carlo experiments are rerun and the amended results reported.

The rest of the paper is organized as follows. Section 2 summarizes the general approach to bias correction. Section 3 presents the main theoretical results corresponding to the SAR model. Section 4 presents Monte Carlo results. Section 5 concludes.

2 A General Method for Bias-Correction

In this section, we lay out a general framework for bias and variance corrections, which is applicable to a class of models (including the SAR model) containing one nonlinear parameter (bias incurring parameter) and some other linear (location) and scale parameters. We first present third-order stochastic expansions based on a concentrated estimating equation for a nonlinear parameter, and then we outline the main ideas of bootstrap method for estimating quantities in the third-order bias and variance formulas.

2.1 Third-Order Bias and Variance of a Nonlinear Estimator

Bao and Ullah (2007a) considered a general class of $\sqrt{n}$-consistent estimators identified by the moment condition or estimating equation

$$\hat{\theta}_n = \arg\{\psi_n(\theta) = 0\} \tag{6}$$

where $\psi_n(\theta) \equiv \psi_n(Z_n; \theta)$ is a $k \times 1$ vector-valued function of the observable data $Z_n = \{Z_i\}_{i=1}^n$, iid or non-iid, and a parameter vector θ. Furthermore, $\psi_n(\theta)$ is of the same dimension as θ and is normalized to have order $O_p(n^{-1/2})$.[5] This framework extends that of Rilstone et al. (1996) who considered $\psi_n(\theta) = \frac{1}{n}\sum_{i=1}^n q_i(\theta_n)$ where the summands are assumed to be iid. They obtain a third-order expansion for $\hat{\theta}_n$, and a second-order bias and a third-order MSE for $\hat{\theta}_n$.

All the results discussed earlier build upon the assumption that $\mathrm{E}\psi_n(\theta) = 0$, a common assumption for a consistent joint estimation. This is usually true if θ contains all the

[5] This is in fact a generalized version of the well-known M-estimation (maximum likelihood type estimation) of Huber (1964). Obviously, the maximum likelihood or quasi-maximum likelihood, least squares, method of moments, and generalized method of moments are the special cases of this estimation method.

model parameters, e.g., $\psi_n(\theta)$ is the full score function. Occasionally, this is also true when concentrated score function is used, but with stronger conditions, such as normality, imposed; see, e.g., the pure SAR model with σ^2 unknown (Bao and Ullah, 2007a). In general, however, this condition is violated if the moment condition is associated with the concentrated estimating function where the parameters not of direct interest are replaced by their estimators, often constrained upon the parameter of interest. For example, in the SAR model with regressors considered in this paper, we are only interested in λ, the spatial lag parameter, but the regression coefficients β and the error variance σ^2 are present.[6]

As pointed out in the introduction, it is very much desirable to construct a stochastic expansion based on concentrated estimating equation. First, it is much easier to work with a scalar case than with a vector to derive the second-order results (Bao and Ullah, 2007a, p.400). Second, as is well known, an ML or QML estimation process often involves two sets of parameters, where the QMLE of the first set (e.g., pure mean and scale parameters) given the second has an explicit expression. Thus, the likelihood can be 'concentrated' by replacing the first set of parameters by their constrained QMLEs. As a result, the optimization procedure can be greatly simplified. The third and perhaps the strongest reason for working with the concentrated estimating equation is that often the estimation of some parameter(s) in the model incurs bias but the estimation of other parameters alone doesn't or incurs very little bias. For example for the SAR model considered in this paper the estimation of λ incurs bias but the estimation of β and σ (given λ) doesn't or incurs very little bias. In a Monte Carlo study in comparing the method of moment estimators, GMM estimators, and QMLEs for the SAR model, Lee (2007a) noted that the main differences of various estimation approaches are on the estimation of the spatial effect λ. Thus, it is only necessary to focus on the QMLE $\hat{\lambda}_n$ of λ for bias-correction. Once a bias-corrected estimator of λ, $\hat{\lambda}_n^{\text{bc}}$ say, is given, the resulted estimators for β and σ^2 obtained by plugging $\hat{\lambda}_n^{\text{bc}}$ into (3) and (4) can be expected be nearly unbiased.

To fix the idea, let $\theta = (\lambda', \alpha')'$ where λ is the scalar nonlinear parameter of which the estimation incurs biasness, and given λ the estimation of the parameter vector α has an analytical solution. Let $\hat{\alpha}_n(\lambda)$ be the constrained estimator of α for a given λ value. Let $\theta_0 = (\lambda_0', \alpha_0')'$ be the true value of the parameter vector. Partition $\psi_n(\theta)$ according

[6]Making inference about the parameter of interest in the presence of many parameters not of direct interest (called the nuisance parameters) is a standard statistical problem, and it is typical in these situations to replace the nuisance parameters by their estimators in the object function or the estimating function. There is a vast literature on the satisfactory handling of nuisance parameters. Most of this work has focused on the modification of the likelihood function and the concentrated likelihood function. See Laskar and King (1998) for a survey and a comparison of the various methods.

to $(\lambda', \alpha')'$, i.e., $\psi_n(\theta) = \{\psi'_{\lambda n}(\lambda, \alpha), \psi'_{\lambda n}(\lambda, \alpha)\}'$. Define $\tilde{\psi}_n(\lambda) \equiv \psi_{\lambda n}(\lambda, \hat{\alpha}(\lambda))$, called the *concentrated estimating equation* (CEE). Then, the estimator $\hat{\lambda}_n$ of λ would typically be

$$\hat{\lambda}_n = \arg\{\tilde{\psi}_n(\lambda) = 0\}. \tag{7}$$

The CEE given above looks identical to the *joint estimating equation* (JEE) considered in Rilstone et al. (1996) and Bao and Ullah (2007a, b). Thus, one would expect that the stochastic expansion for $\hat{\lambda}_n$ takes the same form, though the regularity conditions need to be strengthened. However, there is a major difference: the expectation of $\tilde{\psi}_n(\lambda_0)$ may not be zero even if $\mathrm{E}\psi_n(\theta_0) = 0$. Thus, blindly applying the formulas derived under JEE may lead to misleading results. If $\hat{\alpha}_n(\lambda)$ is $\sqrt{n}$-consistent, it is typical that $\mathrm{E}[\tilde{\psi}_n(\lambda_0)] = O(n^{-1})$, i.e., the expectation goes to zero at an n-rate. If this is true, then $\mathrm{E}[\tilde{\psi}_n(\lambda_0)]$ constitutes an important term in the bias correction. In this case, the bias formula need to be modified. As a consequence, the higher-order approximations to the variance needs to be modified as well. The mean squared error (MSE), however, remains in the same form as it directly follows the stochastic expansions for $\hat{\lambda}_n$.

Let $H_{rn}(\lambda) = d^r\tilde{\psi}(\lambda)/d\lambda^r, r = 1, 2, 3$. Let $\tilde{\psi}_n \equiv \tilde{\psi}_n(\lambda_0)$, $H_{rn} \equiv H_{rn}(\lambda_0)$ and $H^\circ_{rn} = H_{rn} - \mathrm{E}H_{rn}, r = 1, 2, 3$. Define $\Omega_n = -E(H_{1n})^{-1}$. Note that here and hereafter the expectation operator corresponds to the true model or the true parameter values θ_0. Let Λ be the parameter space of λ. So far we have not yet specified the form of the $\tilde{\psi}_n(\lambda)$ function, thus as general theories we need some generic smoothness conditions on $\tilde{\psi}_n(\lambda)$, as those of Bao and Ullah (2007a) for a JEE. We feel, however, the regularity conditions of Bao and Ullah (2007a) need to be tightened under the CEE, which are given below.

Assumption A. *Λ is compact with λ_0 being an interior point. $E(\tilde{\psi}_n) = O(n^{-1})$, and $\hat{\lambda}_n$, as a solution of $\tilde{\psi}_n(\lambda) = 0$, is a $\sqrt{n}$-consistent estimator of λ_0.*

Assumption B. *$\tilde{\psi}_n(\lambda)$ is differentiable up to rth order for λ in a neighborhood of λ_0,* $\mathrm{E}(H_{rn}) = O(1)$, and $H^\circ_{rn} = O_p(n^{-\frac{1}{2}}), r = 1, 2, 3$.

Assumption C. $\mathrm{E}(H_{1n})^{-1} = O(1)$, *and $H_{1n}^{-1} = O_p(1)$.*

Assumption D. $|H_{rn}(\lambda) - H_{rn}(\lambda_0)| \leq |\lambda - \lambda_0|U_n$ *for λ in a neighborhood of $\lambda_0, r = 1, 2, 3$, and $E(|U_n|) < C < \infty$ for some constant C.*

The $\sqrt{n}$-consistency is a standard requirement for a higher-order stochastic expansion. In the context of CEE, the $\sqrt{n}$-consistency of $\hat{\lambda}_n$ implies $\mathrm{E}(\tilde{\psi}_n) = o(n^{-1/2})$ but not zero in general due to the estimation of the nuisance parameters. If the estimators of the nuisance parameters are also $\sqrt{n}$-consistent, it can be argued that $\mathrm{E}(\tilde{\psi}_n) = O(n^{-1})$. Further, the $\sqrt{n}$-consistency of $\hat{\lambda}_n$ implies $\tilde{\psi}_n = O_p(n^{-\frac{1}{2}})$. The Assumptions B and C are the tightened versions of the Assumptions 4 and 5 in Bao and Ullah (2007a). $\mathrm{E}(H_{rn}) = O(1)$ and $H_{rn}^{\circ} = O_p(n^{-\frac{1}{2}})$ are needed so that H_{rn} can be replaced by $\mathrm{E}(H_{rn})$ at an appropriate place in the expansion with the error $O_p(n^{-\frac{1}{2}})$ being absorbed into the overall error term.[7] We are ready to state the general theorems. All the proofs are given in Appendix A.

Theorem 2.1. *Let Assumptions A-D hold for some $r \geq 3$. Then, a third-order stochastic expansion for $\hat{\lambda}_n$ is give by*

$$\hat{\lambda}_n - \lambda_0 = a_{-1/2} + a_{-1} + a_{-3/2} + O_p(n^{-2}), \tag{8}$$

where $a_{-s/2}$ represents terms of order $O_p(n^{-s/2})$ for $s = 1, 2, 3$, and they are $a_{-1/2} = \Omega_n \tilde{\psi}_n$, $a_{-1} = \Omega_n H_{1n}^{\circ} a_{-1/2} + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2}^2)$, and $a_{-3/2} = \Omega_n H_{1n}^{\circ} a_{-1} + \frac{1}{2}\Omega_n H_{2n}^{\circ}(a_{-1/2}^2) + \Omega_n \mathrm{E}(H_{2n})(a_{-1/2}a_{-1}) + \frac{1}{6}\Omega_n \mathrm{E}(H_{3n})(a_{-1/2}^3)$.

The third-order expansion based on CEE is seen to have an identical form as those in Rilstone et al. (1996) under a special JEE with iid summands, in Bao and Ullah (2007b) under a special JEE with non-iid summands, and in Bao and Ullah (2007a) under a general JEE. Our set-up is more closely related to Bao and Ullah (2007a) by concentrating their JEE to give our CEE. However, Bao and Ullah (2007a) did not give a detailed proof of the results, and there seems to be a need for a more detailed regularity conditions for the expansions to hold with proper orders. Based on this third-order stochastic expansion for the nonlinear estimator $\hat{\lambda}_n$, we can easily obtain second-order and third-order expansions for the bias, the MSE, and the variance of $\hat{\lambda}_n$.

Corollary 2.1. *Under the assumptions of Theorem 1, assume further that a quantity bounded in probability has a finite expectation. Then, we have a third-order expansion for the bias of $\hat{\lambda}_n$,*

$$\mathrm{Bias}(\hat{\lambda}_n) = b_{-1} + b_{-3/2} + O(n^{-2}), \tag{9}$$

where $b_{-s/2} = O(n^{-s/2}), s = 2, 3$, with $b_{-1} = \mathrm{E}(a_{-1/2} + a_{-1})$ and $b_{-3/2} = \mathrm{E}(a_{-3/2})$.

Note that $\mathrm{E}(a_{-1/2}) = \Omega_n \mathrm{E}(\tilde{\psi}_n)$. This term is $O(n^{-1})$ under CEE, and is identically zero when JEE is used. Rilstone et al. (1996) and Bao and Ullah (2007a, b) considered only second-order expansions for the bias. Their formulas correspond to our b_{-1} term only. Comparing with their expansions for the bias, we see that b_{-1} contains an extra term,

[7] Under a specific model and a specific estimation method (such as the SAR model estimated by the QML method), the form of the $\tilde{\psi}_n(\lambda)$ function is known, and these generic conditions are satisfied under a set of weak and primitive conditions. The Assumption A may be relaxed to allow for asymptotic (first-order) bias, and our methods can in principle be applied to do higher-order bias reduction for dynamic or nonlinear panel models with fixed effects, see Hahn and Kuersteiner (2002) and Hahn and Newey (2004).

$2\Omega_n \mathrm{E}(\tilde{\psi}_n)$. When CEE is used, this term plays a key role, which means that blindly using the formula of Bao and Ullah (2007a) can give misleading results. This point is confirmed by the Monte Carlo results presented in Section 4.

Adding a third-order bias-correction term $b_{-3/2}$ into the formula gives us a choice for further improvement on the performance of the bias-correction procedure if necessary. With the results of Corollary 2.1, a second-order bias-corrected estimator of λ is $\hat{\lambda}_n^{\mathrm{bc2}} = \hat{\lambda}_n - b_{-1}$, and a thir-order bias-corrected estimator is $\hat{\lambda}_n^{\mathrm{bc3}} = \hat{\lambda}_n - b_{-1} - b_{-3/2}$. However, the practical implementation of the third-order bias-correction requires the estimation of $b_{-3/2}$, which is surely complicated or even prohibitive if the analytical approach is followed, but adds only a little computationally if the bootstrap procedure introduced in this paper is followed. Similarly, one has a third-order expansion for the MSE of $\hat{\lambda}_n$.

Corollary 2.2. *Under the assumptions of Theorem 1, assume further that a quantity bounded in probability has a finite expectation. Then, we have a third-order expansion for the MSE of* $\hat{\lambda}_n$*,*

$$\mathrm{MSE}(\hat{\lambda}_n) = m_{-1} + m_{-3/2} + m_{-2} + O(n^{-5/2}), \tag{10}$$

where $m_{-s/2} = O(n^{-s/2}), s = 2, 3, 4,$ *with* $m_{-1} = \mathrm{E}(a_{-1/2}^2)$*,* $m_{-3/2} = 2\mathrm{E}(a_{-1/2}a_{-1})$*, and* $m_{-2} = 2\mathrm{E}(a_{-1/2}a_{-3/2} + a_{-1}^2)$*.*

Clearly, the leading term $m_{-1} = \Omega_n^2 \mathrm{E}(\tilde{\psi}_n^2)$ in the third-order expansion for $\mathrm{MSE}(\hat{\lambda}_n)$ is the asymptotic variance of $\hat{\lambda}$, $m_{-1} + m_{-3/2}$ gives a second-order expansion, and $m_{-1} + m_{-3/2} + m_{-2}$ gives a third-order expansion for $\mathrm{MSE}(\hat{\lambda}_n)$.[8]

While it is important to have higher-order expansions for $\mathrm{MSE}(\hat{\lambda}_n)$ for the purpose of efficiency comparison, it is more important to have higher-order expansions for the variance of $\hat{\lambda}_n$ for inference purpose. In doing so, one is tempted to simply combine the above expansions for the bias and MSE to give second- and third-order expansions:

$$\mathrm{Var}(\hat{\lambda}_n) = m_{-1} + m_{-3/2} + O(n^{-2}), \tag{11}$$

$$\mathrm{Var}(\hat{\lambda}_n) = m_{-1} + m_{-3/2} + m_{-2} + b_{-1}b_{-1}' + O(n^{-5/2}). \tag{12}$$

Theoretically these are correct, but empirically they do not guarantee positivity of the variance estimator when n is not large (see Section 4.1 for more discussions on this). We thus propose to have variance expansions directly out of the stochastic expansions for $\hat{\lambda}_n$.

[8]Rilstone et al. (1996) and Bao and Ullah (2007a, b) refer to the expansion of the form in (10) as a second-order expansion, but it should in fact be a third-order expansion. Under the special JEE considered by Rilstone et al. (1996), it can easily be shown that $\mathrm{E}(a_{-1/2}a_{-1}) = O(n^{-2})$, the supposedly second-order term becomes a part of the third-order term m_{-2}. However, in general $\mathrm{E}(a_{-1/2}a_{-1}) = O(n^{-3/2})$.

Corollary 2.3. *Under the assumptions of Theorem 1, assume further that a quantity bounded in probability has a finite expectation. Then, we have a third-order expansion for the variance of* $\hat{\lambda}_n$*,*

$$\mathrm{Var}(\hat{\lambda}_n) = v_{-1} + v_{-3/2} + v_{-2} + O(n^{-5/2}), \tag{13}$$

where $v_{-s/2} = O(n^{-s/2}), s = 2, 3, 4,$ *and are of the forms,* $v_{-1} = \mathrm{Var}(a_{-1/2})$*,* $v_{-1} + v_{-3/2} = \mathrm{Var}(a_{-1/2} + a_{-1})$ *and* $v_{-1} + v_{-3/2} + v_{-2} = \mathrm{Var}(a_{-1/2} + a_{-1} + a_{-3/2})$*.*

With the results of Corollaries 2.1 and 2.3, one can correct $\hat{\lambda}_n$ and its standard error (se) for an improved inference for λ. Clearly, (13) reduces to (12) after dropping the terms of the same order as the remainder. The expressions for the terms in all the second expansions given above are fairly short and simple, but the expressions for the third-order term in all the expansions are long but straightforward to obtain. We will present these in the framework of SAR model in the next section. The third-order expansions are presented by clearly separating out the terms of different order, thus allowing one to choose between the 2nd- or 3rd-order approximations according to the actual needs. For example, if one feels that second-order approximations are sufficient, then one simply drops the third-order terms $a_{-3/2}, b_{-3/2}, m_{-2}$, and v_{-2} in the expansions, which results in much simpler expressions.

2.2 A bootstrap method for estimating the bias and variance corrections

The second- or third-order corrections on the bias and variance of nonlinear estimators are practically tractable only if one could find a simple way to estimate the quantities like $\mathrm{E}(H_{rn})$, $\mathrm{E}(\tilde{\psi}_n^2)$, $\mathrm{E}(H_{1n}\tilde{\psi}_n^2)$, etc. The analytical approach is to first find these expectations and then replace θ in the resulted expressions by its consistent estimator $\hat{\theta}_n$. However, finding these expectations analytically seems to be an impossible task unless one is using a simple model with iid normal errors. We now discuss a simple bootstrap method for estimating these quantities. Consider a general model of the form

$$g(Z_n, \theta_0) = u_n$$

where u_n is the disturbance vector of iid (not necessarily normal) components. Assume that the key quantities $\tilde{\psi}_n$ and H_{rn} can be expressed as $\tilde{\psi}_n \equiv \tilde{\psi}_n(u_n, \theta_0)$ and $H_{rn} \equiv H_{rn}(u_n, \theta_0), r = 1, 2, 3$. Let $\hat{u}_n = g(Z_n, \hat{\theta}_n)$ be the vector of estimated residuals based on the data. Resample the elements of $\hat{u}_n$ (by making n random draws with replacement) to give $\hat{u}_{n,b}$, and compute $\tilde{\psi}_n(u_{n,b}, \hat{\theta}_n)$ and $H_{rn}(u_{n,b}, \hat{\theta}_n)$. Repeat this procedure for $b = 1, 2, \cdots, B$ times, and the bootstrap estimates of these expected quantities are given as, e.g.,

$$\widehat{\mathrm{E}}(\tilde{\psi}_n^k) = \frac{1}{B}\sum_{b=1}^{B} \tilde{\psi}_n^k(u_{n,b}, \hat{\theta}_n), k = 1, 2, \cdots,$$

$$\begin{aligned}
\widehat{\mathrm{E}}(H_{rn}) &= \frac{1}{B}\sum_{b=1}^{B} H_{rn}(u_{n,b}, \hat{\theta}_n), \\
\widehat{\mathrm{E}}(\tilde{\psi}_n^2 H_{rn}^2) &= \frac{1}{B}\sum_{b=1}^{B} \tilde{\psi}_n^2(u_{n,b}, \hat{\theta}_n) H_{rn}^2(u_{n,b}, \hat{\theta}_n), \\
\cdots &= \cdots .
\end{aligned}$$

Plugging these bootstrap estimates into the bias, MSE and variance formulas, we obtain the bootstrap-based estimates of bias, MSE, and variance of $\hat{\lambda}_n$.

Note that in the entire bootstrap process, the same estimate $\hat{\theta}_n$ based on the original data is used when recalculating $\tilde{\psi}_n$ and H_{rn} based on each bootstrap sample $\hat{u}_{n,b}$. The reestimation of the model parameter θ is thus avoided, which makes this bootstrap procedure easy to be implemented. In summary, our proposed approach for bias and se corrections takes the advantages of both stochastic expansions and bootstrap, neither of which alone allows us to handle the problem of this type comfortably. The usefulness and effectiveness of this approach is fully demonstrated in the following sections using the SAR model.

3 Bias-Corrected Estimation for SAR Model

We now consider the estimation of spatial lag parameter λ in the general SAR model specified in (1), to give a detailed demonstration on the applications of the general methods presented in the early section. The nature of the SAR model indeed renders it a special attention in terms of bias and se corrections. First, the parameter λ enters into the model in nonlinear manner, hence it is likely that the estimation of it would incur bias. Second, the degree of spatial dependence among the spatial units depends not only on the magnitude of the spatial parameter λ, but also on the number of neighbors each spatial unit has, or equivalently the number of non-zero elements that each row of the W_n matrix contains. A very important special case of this is that the number of neighbors, h_n say, grows with n (see, e.g., Case, 1991), and in this case, Lee (2004a) showed that the QML estimators of λ and β are no longer $\sqrt{n}$-consistent, but rather $\sqrt{n/h_n}$-consistent. Thus, the effective sample size is n/h_n, and the bias and variance formulas given above need to be adjusted to allow for this possibility. Conceptually, this may be fairly straightforward as one may simply replace n by h_n/n everywhere in the expansion formulas. Theoretically, however, much needs to be done in terms of regularity conditions and formal proofs of the results. We do so in this paper by following the theoretical foundations laid out in Lee (2004a). Another challenging problem is that once the general formulas for second-order bias and variance are available, how do we implement them for practical applications? Below are the main theoretical results, followed by a simple bootstrap procedure for estimating the expectations of various ratios of quadratic forms in the bias and variance corrections.

3.1 The main results

Continuing on the QML estimation of the SAR model outlined in the introduction, we see that the QMLE $\hat{\lambda}_n$ of the spatial parameter λ, which maximizes the concentrated log-likelihood function $\ell_n^c(\lambda)$ given in (5), can be equivalently defined as

$$\hat{\lambda}_n = \arg\{\tilde{\psi}_n(\lambda) = 0\},$$

where $\tilde{\psi}_n(\lambda)$ is the derivative of $\frac{h_n}{n}\ell_n^c(\lambda)$ and has the form,

$$\tilde{\psi}_n(\lambda) = -h_n T_{0n}(\lambda) + h_n R_{1n}(\lambda), \tag{14}$$

where $T_{0n}(\lambda) = n^{-1}\text{tr}(G_n(\lambda))$, $G_n(\lambda) = W_n A_n^{-1}(\lambda)$, and

$$R_{1n}(\lambda) = \frac{Y_n' A_n'(\lambda) M_n W_n Y_n}{Y_n' A_n'(\lambda) M_n A_n(\lambda) Y_n}.$$

Clearly the function $\tilde{\psi}_n(\lambda)$ defined in (14) leads to a concentrated estimating function (CEE), and fits into the general framework described in Section 2. However, there is a difference – the quantity h_n measuring the degree of spatial dependence may alter the rate of convergence of the QMLEs in the first-order asymptotics (Lee, 2004a) and of course the magnitude of the quantities in the higher-order asymptotics. Paralleled with the general theories given in Section 2. we now present a complete and rigorous treatment for the SAR model, taking into account of the possibility that h_n is unbounded.

The rth derivative of $\tilde{\psi}_n(\lambda)$, $H_{rn}(\lambda) = \frac{d^r}{d\lambda^r}\tilde{\psi}_n(\lambda), r = 1, 2, 3$, are

$$h_n^{-1} H_{1n}(\lambda) = -T_{1n}(\lambda) - R_{2n}(\lambda) + 2R_{1n}^2(\lambda), \tag{15}$$

$$h_n^{-1} H_{2n}(\lambda) = -2T_{2n}(\lambda) + 2R_{1n}(\lambda)R_{2n}(\lambda) + 8R_{1n}^3(\lambda), \tag{16}$$

$$h_n^{-1} H_{3n}(\lambda) = -6T_{3n}(\lambda) + 2R_{2n}^2(\lambda) + 10R_{1n}^2(\lambda)R_{2n}(\lambda) - 32R_{1n}^4(\lambda), \tag{17}$$

where $T_{rn}(\lambda) = n^{-1}\text{tr}(G_n^{r+1}(\lambda)), r = 0, 1, 2, 3$, and

$$R_{2n}(\lambda) = \frac{Y_n' W_n' M_n W_n Y_n}{Y_n' A_n'(\lambda) M_n A_n(\lambda) Y_n}.$$

Recall $\tilde{\psi}_n = \tilde{\psi}_n(\lambda_0)$ and $H_{rn} = H_{rn}(\lambda_0), r = 1, 2, 3$. Similarly, let $A_n = A_n(\lambda_0)$, $G_n = G_n(\lambda_0)$, $T_{rn} = T_{rn}(\lambda_0), r = 0, 1, 2, 3$, and $R_{rn} = R_{rn}(\lambda_0), r = 1, 2$. Let $\eta_n = G_n X_n \beta_0$.

The regularity conditions listed in Section 2 are generic as the model and the estimation method had not yet specified. With the specification of the SAR model and the quasi-maximum likelihood estimation method, the $\tilde{\psi}_n(\lambda)$ function and its derivatives are completed known. Thus, the regularity conditions can be made more specific or more primitive. First, the set of rather primitive conditions of Lee (2004a) for the $\sqrt{n/h_n}$-consistency of the QMLE $\hat{\lambda}_n$ are essential and are summarized below in Assumptions 1-6.

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

Assumption 2: *The innovations* $\{u_{n,i}\}$ *are iid with mean zero and variance* σ^2. $\mathrm{E}|u_{n,i}|^{4+\gamma}$ *exists for some* $\gamma > 0$.

Assumption 3: *The elements* $w_{n,ij}$ *of* W_n *are at most of order* h_n^{-1} *uniformly for all* i *and* j, *where the rate sequence* $\{h_n\}$ *can be bounded or divergent but satisfying* $h_n^{1+\epsilon}/n \to 0$ *for some* $\epsilon > 0$ *as* $n \to \infty$. *As a normalization,* $w_{n,ii} = 0$ *for all* i. *Furthermore, the matrices* $\{W_n\}$ *are uniformly bounded in both row and column sums.*

Assumption 4: *The matrix* A_n *is nonsingular,* $\{A_n^{-1}\}$ *are uniformly bounded in both row and column sums, and* $\{A_n^{-1}(\lambda)\}$ *are uniformly bounded in either row or column sum, uniformly in* $\lambda \in \Lambda$.

Assumption 5: *The elements of the* $n \times k$ *matrix* X_n *are uniformly bounded for all* n, *and* $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ *exists and is nonsingular.*

Assumption 6: *The elements of* $M_n \eta_n$ *have the uniform order* $O(1/\sqrt{h_n})$, *and*

$$\lim_{n\to\infty} \frac{h_n}{n} \eta_n' M_n \eta_n = c$$

where either $c > 0$; *or* $c = 0$ *but* $\lim_{n\to\infty} \frac{h_n}{n}(\ln|\sigma_0^2 A_n^{-1} A_n'^{-1}| - \ln|\sigma_n^2(\lambda) A(\lambda)_n^{-1} A'(\lambda)_n^{-1}|) \neq 0$, *whenever* $\lambda \neq \lambda_0$, *where* $\sigma_n^2(\lambda) = \frac{\sigma_0^2}{n} \mathrm{tr}[A_n'^{-1} A'(\lambda) A(\lambda) A_n^{-1}]$.

The Assumptions 1-6 listed above are the Assumptions 1, 2, 3′, 4-7, and 10 of Lee (2004a). Under these assumptions, the QMLE $\hat{\lambda}_n$ is a $\sqrt{n/h_n}$-consistent estimator of λ_0. In the regular case where h_n is bounded, i.e., the degree of spatial dependence does not grow with the sample size, $\hat{\lambda}_n$ becomes $\sqrt{n}$-consistent. These assumptions lead to the first-order asymptotics for $\hat{\lambda}_n$, which are shown to be essential as well for the third-order stochastic expansion for $\hat{\lambda}_n$, and the third-order expansions for the bias, MSE, and variance of $\hat{\lambda}_n$. Some further conditions are needed for ensuring proper orders of R_{1n} and R_{2n}, which are crucial for the proper behaviors of the derivatives $H_{rn}, r = 1, 2, 3$.

Assumption 7: *(i)* $\mathrm{E}[\frac{h_n}{n}(Y_n' A_n' M_n W_n Y_n)(\bar{\sigma}_n^{-4} - \sigma_0^{-4})(\hat{\sigma}_n^2 - \sigma_0^2)] = O(n^{-1})$; *and (ii)* $\mathrm{E}[\frac{h_n}{n}(Y_n' W_n' M_n W_n Y_n)(\bar{\sigma}_n^{-4} - \sigma_0^{-4})(\hat{\sigma}_n^2 - \sigma_0^2)] = O(n^{-1})$, *where* $\bar{\sigma}_n^2$ *lies between* $\hat{\sigma}_n^2$ *and* σ_0^2.

These conditions are reasonable as under the early assumptions $\frac{h_n}{n}(Y_n' A_n' M_n W_n Y_n) =$

$O_p(1)$, $\frac{h_n}{n}(Y_n'W_n'M_nW_nY_n) = O_p(1)$, $\hat{\sigma}_n^2 - \sigma_0^2 = O_p(n^{-\frac{1}{2}})$, and $\bar{\sigma}_n^{-4} - \sigma_0^{-4} = O_p(n^{-\frac{1}{2}})$. Thus, the two quantities inside the expectation sign are both $O_p(n^{-1})$. To ensure the proper stochastic behavior of $H_{rn}, r = 1, 2, 3$, the following conditions are needed.

Assumption 8: (i) $h_n^s \mathrm{E}[(R_{1n} - \mathrm{E}R_{1n})^s] = O((\frac{h_n}{n})^{\frac{1}{2}}), s = 2, 3, 4$; (ii) $h_n^2 \mathrm{E}[(R_{2n} - \mathrm{E}R_{2n})^2] = O((\frac{h_n}{n})^{\frac{1}{2}})$; and (iii) $h_n^{s+1}\mathrm{E}[(R_{1n} - \mathrm{E}R_{1n})^s(R_{2n} - \mathrm{E}R_{2n})] = O((\frac{h_n}{n})^{\frac{1}{2}}), s = 1, 2$.

The three conditions in Assumption 8 are in fact rather weak, since following the results of Lemma 3.1 below, all the random quantities inside the expectation sign are of order $O_p(\frac{h_n}{n})$ or lower, and hence their expectations are likely stay with the same order but nonstochastic, i.e., $O(\frac{h_n}{n})$ or lower. However, what is needed is only that the expectation of those quantities are of order $O((\frac{h_n}{n})^{\frac{1}{2}})$. We have the following important lemma.

Lemma 3.1. *Under the Assumptions 1-7, we have (i)* $h_nR_{1n} = \mathrm{E}(h_nR_{1n}) + O_p((\frac{h_n}{n})^{\frac{1}{2}})$, *and (ii)* $h_nR_{2n} = \mathrm{E}(h_nR_{2n}) + O_p((\frac{h_n}{n})^{\frac{1}{2}})$.

The proof of this lemma is given in Appendix B. Using the results of this lemma and under the conditions listed above, we are able to prove the following theorem and corollaries. These theorem and corollaries parallel the general theorem and corollaries given in Section 2 with the order of magnitude of each term adjusted to allow for the possibility that h_n increases with n. Their proofs are given in Appendix B.

Theorem 3.1. *Under the Assumptions 1-8, we have a third-order stochastic expansion,*

$$\hat{\lambda}_n - \lambda_0 = a_{-1/2} + a_{-1} + a_{-3/2} + O_p((n/h_n)^{-2}), \tag{18}$$

where $a_{-s/2} = O_p((n/h_n)^{-s/2}), s = 1, 2, 3$, *and are of the form:* $a_{-1/2} = \Omega_n\tilde{\psi}_n$,

$$\begin{aligned} a_{-1} &= \Omega_n\tilde{\psi}_n + \Omega_n^2 H_{1n}\tilde{\psi}_n + \frac{1}{2}\Omega_n^3 \mathrm{E}(H_{2n})\tilde{\psi}_n^2, \text{ and} \\ a_{-3/2} &= \Omega_n\tilde{\psi}_n + 2\Omega_n^2 H_{1n}\tilde{\psi}_n + \Omega_n^3\mathrm{E}(H_{2n})\tilde{\psi}_n^2 + \Omega_n^3 H_{1n}^2\tilde{\psi}_n + \frac{1}{2}\Omega_n^3 H_{2n}\tilde{\psi}_n^2 \\ &\quad + \frac{3}{2}\Omega_n^4\mathrm{E}(H_{2n})H_{1n}\tilde{\psi}_n^2 + \frac{1}{2}\Omega_n^5\mathrm{E}(H_{2n})^2\tilde{\psi}_n^3 + \frac{1}{6}\Omega_n^4\mathrm{E}(H_{3n})\tilde{\psi}_n^3, \end{aligned}$$

where $\tilde{\psi}_n$ *and* $H_{rn}, r = 1, 2, 3$ *are given in (14)-(17), and* $\Omega_n = -\mathrm{E}(H_{1n})^{-1}$.

Note that $a_{-s/2}, s = 1, 2, 3$, in fact have the same expressions as those in Theorem 2.1. The difference is in their stochastic orders. By applying this stochastic expansion, one obtains immediately the expansions for the bias, the MSE, and the variance of $\hat{\lambda}_n$.

Corollary 3.1. *Let the Assumptions 1-8 hold. If* $E(R_{1n}R_{2n})$ *and* $E(R_{1n}^3)$ *exist, we have a second-order expansion for the bias of* $\hat{\lambda}_n$,

$$\mathrm{B}_2(\hat{\lambda}_n) = 2\Omega_n\mathrm{E}(\tilde{\psi}_n) + \Omega_n^2\mathrm{E}(H_{1n}\tilde{\psi}_n) + \frac{1}{2}\Omega_n^3\mathrm{E}(H_{2n})\mathrm{E}(\tilde{\psi}_n^2) + O((n/h_n)^{-3/2}); \tag{19}$$

and if $E(R_{2n}^2)$ and $E(R_{1n}^6)$ exist, we have a third-order expansion for the bias of $\hat{\lambda}_n$,

$$\begin{aligned} \mathrm{B}_3(\hat{\lambda}_n) &= 3\Omega_n \mathrm{E}(\tilde{\psi}_n) + 3\Omega_n^2 \mathrm{E}(H_{1n}\tilde{\psi}_n) + \frac{3}{2}\Omega_n^3 \mathrm{E}(H_{2n})\mathrm{E}(\tilde{\psi}_n^2) + \Omega_n^3 \mathrm{E}(H_{1n}^2\tilde{\psi}_n) \\ &\quad + \frac{1}{2}\Omega_n^3 \mathrm{E}(H_{2n}\tilde{\psi}_n^2) + \frac{3}{2}\Omega_n^4 \mathrm{E}(H_{2n})\mathrm{E}(H_{1n}\tilde{\psi}_n^2) + \frac{1}{2}\Omega_n^5 \mathrm{E}(H_{2n})^2 \mathrm{E}(\tilde{\psi}_n^3) \\ &\quad + \frac{1}{6}\Omega_n^4 \mathrm{E}(H_{3n})\mathrm{E}(\tilde{\psi}_n^3) + O((n/h_n)^{-2}). \end{aligned} \tag{20}$$

Note that $\mathrm{B}_2(\hat{\lambda}_n) = \mathrm{E}(a_{-1/2}) + \mathrm{E}(a_{-1}) + O((n/h_n)^{-3/2})$, and $\mathrm{B}_3(\hat{\lambda}_n) = \mathrm{E}(a_{-1/2}) + \mathrm{E}(a_{-1}) + \mathrm{E}(a_{-3/2}) + O((n/h_n)^{-2})$, where $\mathrm{E}(a_{-1/2}) = O((n/h_n)^{-1})$ as shown in Appendix B, and has the same order as $\mathrm{E}(a_{-1})$. This makes sense intuitively as the estimator is consistent and hence there is no first-order bias correction. The result of Corollary 3.1 leads immediately to a second-order bias-corrected QMLE of λ as follows:

$$\hat{\lambda}_n^{\mathrm{bc2}} = \hat{\lambda}_n - 2\widehat{\Omega}_n \widehat{\mathrm{E}}(\tilde{\psi}_n) - \widehat{\Omega}_n^2 \widehat{\mathrm{E}}(H_{1n}\tilde{\psi}_n) - \frac{1}{2}\widehat{\Omega}_n^3 \widehat{\mathrm{E}}(H_{2n})\widehat{\mathrm{E}}(\tilde{\psi}_n^2), \tag{21}$$

where the quantity with an $\hat{}$ denotes the estimate of the corresponding quantity, and the way this estimate is obtained will be discussed in next subsection. This estimator gives a second-order correction. In contrast, if $\mathrm{E}(\tilde{\psi}_n)$ is treated as zero as in the case of joint estimation, then a bias-adjusted estimator is given as,

$$\hat{\lambda}_n^{\mathrm{ba2}} = \hat{\lambda}_n - \widehat{\Omega}_n^2 \widehat{\mathrm{E}}(H_{1n}\tilde{\psi}_n) - \frac{1}{2}\widehat{\Omega}_n^3 \widehat{\mathrm{E}}(H_{2n})\widehat{\mathrm{E}}(\tilde{\psi}_n^2), \tag{22}$$

which misses out a second-order term $2\widehat{\Omega}_n \widehat{\mathrm{E}}(\tilde{\psi}_n)$, and obviously these terms can be very important. It is not difficult to argue that the terms $2\widehat{\Omega}_n \widehat{\mathrm{E}}(\tilde{\psi}_n)$ and $\widehat{\Omega}_n^2 \widehat{\mathrm{E}}(H_{1n}\tilde{\psi}_n)$ are of the same order. This means that the difference between $\hat{\lambda}_n^{\mathrm{bc2}}$ and $\hat{\lambda}_n^{\mathrm{ba2}}$, or between $\hat{\lambda}_n^{\mathrm{bc2}}$ and $\hat{\lambda}_n$, can be quite substantial. Section 4 provides Monte Carlo results for such comparisons.

If the second-order bias-correction is not accurate enough, one can easily go for the third-order correction to have $\hat{\lambda}_n^{\mathrm{bc3}} = \hat{\lambda}_n - \widehat{\mathrm{B}}_3(\hat{\lambda}_n)$, where $\widehat{\mathrm{B}}_3(\hat{\lambda}_n)$ is an estimate of $\mathrm{B}_3(\hat{\lambda}_n)$ given in (20) after dropping the remainder term. Similarly, one obtains the second- and third-order expansions for the MSE of $\hat{\lambda}_n$.

Corollary 3.2. *Let the Assumptions 1-8 hold. If $E(R_{1n}^2 R_{2n})$ and $E(R_{1n}^4)$ exist, then we have a second-order expansion for the MSE of $\hat{\lambda}_n$,*

$$\mathrm{M}_2(\hat{\lambda}_n) = 3\Omega_n^2 \mathrm{E}(\tilde{\psi}_n^2) + 2\Omega_n^3 \mathrm{E}(H_{1n}\tilde{\psi}_n^2) + \Omega_n^4 \mathrm{E}(H_{2n})\mathrm{E}(\tilde{\psi}_n^3) + O((n/h_n)^{-2}); \tag{23}$$

and if $E(R_{2n}^2)$ and $E(R_{1n}^8)$ exist, then we have a third-order expansion for the MSE of $\hat{\lambda}_n$,

$$\begin{aligned} \mathrm{M}_3(\hat{\lambda}_n) &= 6\Omega_n^2 \mathrm{E}(\tilde{\psi}_n^2) + 8\Omega_n^3 \mathrm{E}(H_{1n}\tilde{\psi}_n^2) + 3\Omega_n^4 \mathrm{E}(H_{1n}^2\tilde{\psi}_n^2) + 4\Omega_n^4 \mathrm{E}(H_{2n})\mathrm{E}(\tilde{\psi}_n^3) \\ &\quad + \Omega_n^4 \mathrm{E}(H_{2n}\tilde{\psi}_n^3) + 4\Omega_n^5 \mathrm{E}(H_{2n})\mathrm{E}(H_{1n}\tilde{\psi}_n^3) + \frac{1}{3}\Omega_n^5 \mathrm{E}(H_{3n})\mathrm{E}(\tilde{\psi}_n^4) \\ &\quad + \frac{5}{4}\Omega_n^6 \mathrm{E}(H_{2n})^2 \mathrm{E}(\tilde{\psi}_n^4) + O((n/h_n)^{-5/2}). \end{aligned} \tag{24}$$

As mentioned in Section 2, it is more useful for the purpose of statistical inference to have higher-order expansions for the variances of the original and bias-corrected estimators. It turns out the two agree to order $O(n^{-3/2})$ or $O(n^{-2})$, depending whether a second-order or a third-order expansion is used (see the discussions in Section 3.4).

Corollary 3.3. *Let the Assumptions 1-8 hold. If* $E(R_{1n}^3 R_{2n})$ *and* $E(R_{1n}^5)$ *exist, then we have a second-order expansion for the variance of* $\hat{\lambda}_n$*,*

$$\mathrm{V}_2(\hat{\lambda}_n) = \mathrm{Var}(a_{-1/2} + a_{-1}) + O((n/h_n)^{-2});$$

if $E(R_{2n}^4)$ *and* $E(R_{1n}^{12})$ *exist, then we have a third-order expansion for the variance of* $\hat{\lambda}_n$*,*

$$\mathrm{V}_3(\hat{\lambda}_n) = \mathrm{Var}(a_{-1/2} + a_{-1} + a_{-3/2}) + O((n/h_n)^{-5/2}).$$

For practical applications, we give more explicite expressions for the variance expansions. From Theorem 3.1, $a_{-1/2} + a_{-1} = 2\Omega_n \tilde{\psi}_n + \Omega_n^2 H_{1n} \tilde{\psi}_n + \frac{1}{2}\Omega_n^3 \mathrm{E}(H_{2n})\tilde{\psi}_n^2$, thus,

$$\begin{aligned} & \mathrm{Var}(a_{-1/2} + a_{-1}) \\ = \;& 4\Omega_n^2 \mathrm{Var}(\tilde{\psi}_n) + 4\Omega_n^3 \mathrm{Cov}(\tilde{\psi}_n, H_{1n}\tilde{\psi}_n) + 2\Omega_n^4 \mathrm{E}(H_{2n})\mathrm{Cov}(\tilde{\psi}_n, \tilde{\psi}_n^2) \\ & + \Omega_n^4 \mathrm{Var}(H_{1n}\tilde{\psi}_n) + \Omega_n^5 \mathrm{E}(H_{2n})\mathrm{Cov}(H_{1n}\tilde{\psi}_n, \tilde{\psi}_n^2) + \frac{1}{4}\Omega_n^6 \mathrm{E}(H_{2n})^2 \mathrm{Var}(\tilde{\psi}_n^2). \end{aligned} \quad (25)$$

Again from Theorem 3.1, $a_{-1/2} + a_{-1} + a_{-3/2} = 3\Omega_n \tilde{\psi}_n + 3\Omega_n^2 H_{1n}\tilde{\psi}_n + \frac{3}{2}\Omega_n^3 \mathrm{E}(H_{2n})\tilde{\psi}_n^2 + \Omega_n^3 H_{1n}^2 \tilde{\psi}_n + \frac{1}{2}\Omega_n^3 H_{2n}\tilde{\psi}_n^2 + \frac{3}{2}\Omega_n^4 \mathrm{E}(H_{2n}) H_{1n}\tilde{\psi}_n^2 + \frac{1}{2}\Omega_n^5 \mathrm{E}(H_{2n})^2 \tilde{\psi}_n^3 + \frac{1}{6}\Omega_n^4 \mathrm{E}(H_{3n})\tilde{\psi}_n^3$, thus

$$\begin{aligned} & \mathrm{Var}(a_{-1/2} + a_{-1} + a_{-3/2}) \\ = \;& 9\Omega_n^2 \mathrm{Var}(\tilde{\psi}_n) + 9\Omega_n^4 \mathrm{Var}(H_{1n}\tilde{\psi}_n) + \frac{9}{4}\Omega_n^6 \mathrm{E}(H_{2n})^2 \mathrm{Var}(H_{1n}\tilde{\psi}_n) + \cdots \\ & + 18\Omega_n^3 \mathrm{Cov}(\tilde{\psi}_n, H_{1n}\tilde{\psi}_n) + 9\Omega_n^4 \mathrm{E}(H_{2n})\mathrm{Cov}(\tilde{\psi}_n, \tilde{\psi}_n^2) + \cdots, \end{aligned} \quad (26)$$

where there are all together 8 variance terms and 28 covariance terms.

Clearly, $\mathrm{Var}(a_{-1}) = O((n/h_n)^{-2})$ and thus can be dropped from the second-order variance expansion. Similarly, $\mathrm{Cov}(a_{-1}, a_{-3/2}) = O((n/h_n)^{-5/2})$ and $\mathrm{Var}(a_{-3/2}) = O((n/h_n)^{-3})$, which can be dropped from the third-order variance expansion. These lead to asymptotically equivalent but much simpler variance expansions $\mathrm{V}_2(\hat{\lambda}) = \mathrm{M}_2(\hat{\lambda}) + O((n/h_n)^{-2})$, and $\mathrm{V}_3(\hat{\lambda}) = \mathrm{M}_3(\hat{\lambda}) + \mathrm{B}_2^2(\hat{\lambda}) + O((n/h_n)^{-5/2})$. However, as pointed out in Section 2, these simplifications may not guarantee the positiveness of the variance estimates, thus it is recommanded that the results in (25) and (26) be followed in the practical applications.

3.2 The bootstrap method for practical implementation

From the expressions given in (19) to (26) we see that, in order to calculate various expectations in the bias, MSE, and variance expansions, all we need is to calculate the expectations of R_{1n} and R_{2n} (the ratios of quadratic forms), their powers, and their cross-products of powers, i.e., we need to calculate, for third-order expansions,

$$\mathrm{E}(R_{1n}^k), k = 1, \cdots, 10; \;\; \mathrm{E}(R_{2n}^k), k = 1, \cdots, 4; \;\; \text{and} \;\; \mathrm{E}(R_{1n}^k R_{2n}^m), k = 1, \cdots, 6, m = 1, 2.$$

However, this seems to be an impossible task if the errors are non-normal, and prohibitively complicated if the errors are normal. For the case of a pure SAR model with normal errors, Bao and Ullah (2007a) derived analytical expressions for the expectations of various ratios. However, even for this simple model with normal errors, the expressions are seen to be quite complicated already, and when errors become nonnormal, the analytical expressions are unavailable except the approximations under a small σ (Ullah and Srivastava, 1994).[9]

Thus, for the higher-order results presented above to be practical feasible for a general SAR model, it is highly desirable to have an alternative way to evaluate these expectations. Clearly, it is when the errors are non-normal and the model contains regressors that gives a practical attraction. To solve this puzzle, the bootstrap procedure outlined in Section 2 is made explicite below. Note that the two key quadratic forms can be written as:

$$R_{1n} \equiv R_{1n}(u_n, \theta_0) = \frac{u_n' M_n G_n u_n + u_n' M_n \eta_n}{u_n' M_n u_n},$$
$$R_{2n} \equiv R_{2n}(u_n, \theta_0) = \frac{u_n' G_n' M_n G_n u_n + 2u_n' G_n' M_n \eta_n + \eta_n' M_n \eta_n}{u_n' M_n u_n},$$

where $\eta_n = G_n X_n \beta_0$. These show that $\tilde{\psi}_n = \tilde{\psi}_n(u_n, \theta_0)$, $H_{1n} = H_{1n}(u_n, \theta_0)$, $H_{2n} = H_{2n}(u_n, \theta_0)$, and $H_{3n} = H_{3n}(u_n, \theta_0)$. For the pure SAR process, $\eta_n = 0$, $M_n = I_n$, and these two quantities reduce to $R_{1n} = u_n' G_n u_n / (u_n' u_n)$ and $R_{2n} = u_n' G_n' G_n u_n / (u_n' u_n)$. In other words, all the random quantities in the bias, MSE, and variance formulas can be expressed in terms of u_n and θ_0. This leads naturally to a bootstrap procedure for estimating the expected values of these random quantities (see, e.g., Efron, 1979; Amemiya, 1985, p. 135).

The suggested bootstrap procedure is summarized as follows:

1. Compute the QMLEs $\hat{\theta}_n = (\hat{\beta}_n', \hat{\sigma}_n^2, \hat{\lambda}_n)'$, $\widehat{A}_n$, $\widehat{G}_n$ and $\hat{\eta}_n$,

[9] Working with the joint estimating equation as in Bao (2010) alleviates this problem as only the expectations of quadratic forms, their cross-products and powers are required, but it introduces an additional complication – the expansions become multidimensional involving the parameters β and σ as well, which makes the higher-order corrections on bias and variance extremely difficult. However, one advantage of this analytical approach is that it allows one to see the 'cause' of the bias.

2. Compute QML residuals $\hat{u}_n = \hat{A}_n Y_n - X_n \hat{\beta}_n$, and then center them,

3. Resample $\hat{u}_n$ in the usual way, and denote the resampled vector by $\hat{u}_{n,b}$,

4. Compute $R_{1n}(\hat{u}_{n,b}, \hat{\theta}_n)$ and $R_{2n}(\hat{u}_{n,b}, \hat{\theta}_n)$, and thus all the other quantities $\tilde{\psi}_n(u_{n,b}, \hat{\theta}_n)$, $H_{1n}(u_{n,b}, \hat{\theta}_n)$, $H_{2n}(u_{n,b}, \hat{\theta}_n)$, and $H_{3n}(u_{n,b}, \hat{\theta}_n)$ defined in (13)-(16),

5. Repeat steps (3) and (4) B times to give sequences of bootstrapped values for $\tilde{\psi}_n$, H_{1n}, H_{2n}, and H_{3n}.

The bootstrap estimates of various expectations thus follow. For example, the bootstrap estimates for the means, variances and covariance of $\tilde{\psi}_n^2$ and $H_{1n}\tilde{\psi}_n$ are,

$$
\begin{aligned}
\widehat{\mathrm{E}}(\tilde{\psi}_n^2) &= \frac{1}{B}\sum_{b=1}^{B} \tilde{\psi}_n^2(u_{n,b}, \hat{\theta}_n), \\
\widehat{\mathrm{E}}(H_{1n}\tilde{\psi}_n) &= \frac{1}{B}\sum_{b=1}^{B} H_{1n}(u_{n,b}, \hat{\theta}_n)\tilde{\psi}_n(u_{n,b}, \hat{\theta}_n), \\
\widehat{\mathrm{Var}}(\tilde{\psi}_n^2) &= \frac{1}{B}\sum_{b=1}^{B} \tilde{\psi}_n^4(u_{n,b}, \hat{\theta}_n) - \widehat{\mathrm{E}}(\tilde{\psi}_n^2)^2, \\
\widehat{\mathrm{Var}}(H_{1n}\tilde{\psi}_n) &= \frac{1}{B}\sum_{b=1}^{B} H_{1n}^2(u_{n,b}, \hat{\theta}_n)\tilde{\psi}_n^2(u_{n,b}) - \widehat{\mathrm{E}}(H_{1n}\tilde{\psi}_n)^2, \\
\widehat{\mathrm{Cov}}(\tilde{\psi}_n^2, H_{1n}\tilde{\psi}_n) &= \frac{1}{B}\sum_{b=1}^{B} \tilde{\psi}_n^3(u_{n,b}, \hat{\theta}_n) H_{1n}(u_{n,b}, \hat{\theta}_n) - \widehat{\mathrm{E}}(\tilde{\psi}_n^2)\widehat{\mathrm{E}}(H_{1n}\tilde{\psi}_n).
\end{aligned}
$$

The other quantities in the bias, MSE and variance formulas can be bootstrapped in a similar way. The validity of this bootstrap procedure is supported by the fact that the elements of u_n are iid, which ensures that the elements of $\hat{u}_n$ are asymptotically iid, and the fact that these ratios of quadratic forms are asymptotically equivalent to their counter-parts evaluated at the true parameter value. See, e.g., Efron (1979) and Lahiri (2003) for details on general principles.

3.3 Estimation of the non-spatial parameters

We argued in the introduction that the estimation of the spatial parameter λ is the main source of biasness in the estimation of the general SAR model. Once the QMLE $\hat{\lambda}_n$ of λ is bias-corrected, the estimators of the non-spatial parameters β and σ^2 in the model will be nearly unbiased. We now address this issue in a more formal manner.

Recall $A_n \equiv A_n(\lambda_0) = I_n - \lambda_0 W_n$, $G_n = W_n A_n^{-1}$ and $\eta_n = G_n X_n \beta_0$. We have from (3),

$$\begin{aligned}
\hat{\beta}_n(\hat{\lambda}_n^{\rm bc}) &= (X_n'X_n)^{-1}X_n'A_n(\hat{\lambda}_n^{\rm bc})Y_n \\
&= (X_n'X_n)^{-1}X_n'(I_n - \hat{\lambda}_n^{\rm bc}W_n)Y_n \\
&= (X_n'X_n)^{-1}X_n'[(I_n - \lambda_0 W_n) - (\hat{\lambda}_n^{\rm bc} - \lambda_0)W_n]Y_n \\
&= \hat{\beta}_n(\lambda_0) - (X_n'X_n)^{-1}X_n'W_nY_n(\hat{\lambda}_n^{\rm bc} - \lambda_0) \\
&= \hat{\beta}_n(\lambda_0) - (X_n'X_n)^{-1}X_n'\eta_n(\hat{\lambda}_n^{\rm bc} - \lambda_0) - (X_n'X_n)^{-1}X_n'G_nu_n(\hat{\lambda}_n^{\rm bc} - \lambda_0). \\
{\rm E}[\hat{\beta}_n(\hat{\lambda}_n^{\rm bc})] &= \beta_0 - (X_n'X_n)^{-1}X_n'\eta_n{\rm E}(\hat{\lambda}_n^{\rm bc} - \lambda_0) - (X_n'X_n)^{-1}X_n'G_n{\rm E}[u_n(\hat{\lambda}_n^{\rm bc} - \lambda_0)].
\end{aligned}$$

Now, Assumptions 3 and 4, and Lemma B.1 ensure that G_n is uniformly bounded in both row and column sums. As the elements of X_n are uniformly bounded by Assumption 5, it follows that the elements of η_n are uniformly bounded. Thus, $(X_n'X_n)^{-1}X_n'\eta_n = O(1)$. ${\rm E}(\hat{\lambda}_n^{\rm bc} - \lambda_0)$ is $O((\frac{h_n}{n})^{\frac{3}{2}})$ or $O((\frac{h_n}{n})^2)$ depending on whether $\hat{\lambda}_n^{\rm bc}$ is second- or third-order bias-corrected. Hence, $\hat{\beta}_n(\hat{\lambda}_n^{\rm bc})$ is bias-corrected to order $O((h_n/n)$ or higher if

$$(X_n'X_n)^{-1}X_n'G_n{\rm E}[u_n(\hat{\lambda}_n^{\rm bc} - \lambda_0)] = O((h_n/n)^{\frac{3}{2}}), \ \text{ or smaller.}$$

Although it is difficult to show this in general, it is likely to hold as its stochastic counterpart is $O_p(\sqrt{h_n}/n)$, and the dependence between u_n and $\hat{\lambda}_n^{\rm bc} - \lambda_0$ is likely to be week. Monte Carlo simulation results presented in the next section suggest that it is indeed the case.

The finite sample properties of $\hat{\sigma}_n^2(\hat{\lambda}_n^{\rm bc})$ can be studied in a similar manner. Noting that $Y_n'A_n'M_nA_nY_n = u_n'M_nG_nu_n + u_n'M_n\eta_n$, and that $Y_n'W_n'M_nW_nY_n = u_n'G_n'M_nG_nu_n + u_n'G_n'M_n\eta_n + \eta_n'M_n\eta_n$, we have from (4),

$$\begin{aligned}
\hat{\sigma}_n^2(\hat{\lambda}_n^{\rm bc}) &= \tfrac{1}{n}Y_n'A_n'(\hat{\lambda}_n^{\rm bc})M_nA_n(\hat{\lambda}_n^{\rm bc})Y_n \\
&= \tfrac{1}{n}Y_n'(I_n - \hat{\lambda}_n^{\rm bc}W_n)'M_n(I_n - \hat{\lambda}_n^{\rm bc}W_n)Y_n \\
&= \tfrac{1}{n}Y_n'[I_n - \lambda_0W_n - (\hat{\lambda}_n^{\rm bc} - \lambda_0)W_n)'M_n(I_n - \lambda_0W_n - (\hat{\lambda}_n^{\rm bc} - \lambda_0)W_n]Y_n \\
&= \tfrac{1}{n}Y_n'A_n'M_nA_nY_n - \tfrac{2}{n}Y_n'A_n'M_nW_nY_n(\hat{\lambda}_n^{\rm bc} - \lambda_0) + \tfrac{1}{n}Y_n'W_n'M_nW_nY_n(\hat{\lambda}_n^{\rm bc} - \lambda)^2 \\
&= \hat{\sigma}_n^2(\lambda_0) - \tfrac{2}{n}Y_n'A_n'M_nW_nY_n(\hat{\lambda}_n^{\rm bc} - \lambda_0) + \tfrac{1}{n}Y_n'W_n'M_nW_nY_n(\hat{\lambda}_n^{\rm bc} - \lambda_0)^2 \\
&= \hat{\sigma}_n^2(\lambda_0) - \tfrac{2}{n}{\rm tr}(G_n)(\hat{\lambda}_n^{\rm bc} - \lambda_0) - \tfrac{2}{n}[u_n'M_nG_nu_n + u_n'M_n\eta_n - {\rm tr}(G_n)](\hat{\lambda}_n^{\rm bc} - \lambda_0) \\
&\quad + \tfrac{1}{n}[{\rm tr}(G_nG_n') + \eta_n'M_n\eta_n](\hat{\lambda}_n^{\rm bc} - \lambda_0)^2 + O_p((h_n^{\frac{1}{2}}/n^{\frac{3}{2}})),
\end{aligned}$$

where we have used the results $(\frac{h_n}{n})^{\frac{1}{2}}u_n'M_n\eta_n = O_p(1)$, and $(\frac{h_n}{n})^{\frac{1}{2}}u_n'G_n'M_n\eta_n = O_p(1)$. So, the bias of $\hat{\sigma}_n^2(\hat{\lambda}_n^{\rm bc})$ is of only third-order if

$$\tfrac{1}{n}{\rm E}\left[({\rm tr}(G_nG_n') + \eta_n'M_n\eta_n)(\hat{\lambda}_n^{\rm bc} - \lambda_0)^2 - \tfrac{2}{n}(u_n'M_nG_nu_n + u_n'M_n\eta_n)(\hat{\lambda}_n^{\rm bc} - \lambda_0)\right] = O((\tfrac{h_n}{n})^{\frac{3}{2}}).$$

Again, it is difficult to verify this condition in general. Our simulation results presented in the next section suggest that this term is indeed of a very small magnitude. Note that unlike the constrained QMLE $\hat{\beta}_n^2(\lambda_0)$ which is an unbiased estimator of β_0, the constrained QMLE $\hat{\sigma}_n^2(\lambda_0)$ has to be modified as $\frac{n}{n-k}\hat{\sigma}_n^2(\lambda_0)$ to be an unbiased estimator of σ^2.

3.4 Statistical inference

An immediate gain of the bias-correction and the standard error adjustment on $\hat{\lambda}$ may be the improved finite sample inferences for the model parameters. As we all know, usual inferences for the model parameters are based on the t-ratios. Consider the second-order bias-corrected estimator $\hat{\lambda}_n^{\rm bc2}$ of λ. Denote the second-order bias of $\hat{\lambda}_n$ by $b_2(\hat{\lambda}_n)$. We have from Corollary 3.1, $\mathrm{E}(\hat{\lambda}_n - \lambda_0) = b_2(\hat{\lambda}_n) + O((\frac{h_n}{n})^{3/2})$. Let $\hat{b}_2(\hat{\lambda}_n)$ be a $\sqrt{n/h_n}$-consistent estimator of $b_2(\hat{\lambda}_n)$. As $b_2(\hat{\lambda}_n) = O((\frac{h_n}{n}))$, $\hat{b}_2(\hat{\lambda}_n) = b_2(\hat{\lambda}_n) + O_p((\frac{h_n}{n})^{3/2})$, and $\hat{\lambda}_n^{\rm bc2} = \hat{\lambda}_n - b_2(\hat{\lambda}_n) + O_p((\frac{h_n}{n})^{3/2}) = \lambda_0 + \hat{\lambda}_n - \mathrm{E}(\hat{\lambda}_n) + O_p((\frac{h_n}{n})^{3/2})$. Hence, $M(\hat{\lambda}_n^{\rm bc2}) = \mathrm{E}(\hat{\lambda}_n^{\rm bc2} - \lambda_0)^2 = \mathrm{Var}(\hat{\lambda}_n) + O((\frac{h_n}{n})^2)$. On the other hand, $M(\hat{\lambda}_n^{\rm bc2}) = \mathrm{Var}(\hat{\lambda}_n^{\rm bc2}) + \mathrm{E}(\hat{\lambda}_n^{\rm bc2} - \lambda_0)^2 = \mathrm{Var}(\hat{\lambda}_n^{\rm bc2}) + O((\frac{h_n}{n})^2)$. Thus,

$$\mathrm{Var}(\hat{\lambda}_n^{\rm bc2}) = \mathrm{Var}(\hat{\lambda}_n) + O((h_n/n)^2),$$

i.e., the two variances agree to second-order.

Similarly, if $\hat{\lambda}_n^{\rm bc3}$ is used, the two variances would agree to the third-order in the sense that the difference would be of order $O((\frac{h_n}{n})^{5/2})$. These arguments lead to an important conclusion: to make statistical inferences using a bias-corrected estimator, the standard error (se) of the bias-corrected estimator can simply be taken as the se of the original estimator. Thus, the inference for λ_0 using $\hat{\lambda}_n^{\rm bc2}$ can be carried out based on $V_2(\hat{\lambda}_n)$, and the inference for λ_0 using $\hat{\lambda}_n^{\rm bc3}$ can be carried out based on $V_3(\hat{\lambda}_n)$, given in Corollary 3.3. The implication of this is that the usual t-ratio for inference for λ has a mean distortion,

$$\frac{\hat{\lambda}_n - \lambda_0}{\mathrm{se}(\hat{\lambda}_n)} = \frac{\hat{\lambda}_n - \mathrm{B}(\hat{\lambda}_n) - \lambda_0 + \mathrm{B}(\hat{\lambda}_n)}{\mathrm{se}(\hat{\lambda}_n)} = \frac{\hat{\lambda}_n^{\rm bc} - \lambda_0}{\mathrm{se}(\hat{\lambda}_n^{\rm bc})} + \frac{\mathrm{B}(\hat{\lambda}_n)}{\mathrm{se}(\hat{\lambda}_n)}.$$

Similar arguments apply to the t-ratios for the regression coefficients. Monte Carlo results presented in the next section demonstrates that this mean distortion can be quite big.

4 Monte Carlo Simulation

Extensive Monte Carlo experiments are carried out to investigate (i) the finite sample performance of the QMLE $\hat{\lambda}_n$ and the bias-corrected QMLEs $\hat{\lambda}_n^{\rm bc2}$, $\hat{\lambda}_n^{\rm ba2}$, and $\hat{\lambda}_n^{\rm ba3}$ of the spatial lag parameter λ, and (ii) the impact of the bias and se corrections on $\hat{\lambda}_n$ on the subsequent inferences for λ and other model parameters. Also, through the process of our investigation on the higher-order properties of the SAR model, we found some anomalies in the documented Monte Carlo results in Bao and Ullah (2007a) and in Lee (2004a). Their Monte Carlo experiments are rerun and the amended results reported.

4.1 Monte Carlo Experiment I

To demonstrate the effectiveness of the proposed bias-correction procedure, and to investigate the impact of bias and se corrections on the subsequent interferences, a comprehensive set of Monte Carlo experiments is carried out based on the following general SAR model:

$$Y_n = \lambda W_n Y_n + \beta_0 1_n + X_{n1}\beta_1 + X_{n2}\beta_2 + u_n,$$

where 1_n is an n-vector of ones. For all the Monte Carlo experiments, $\beta = \{\beta_0, \beta_1, \beta_2\}'$ is set at $\{5, 1, 0.5\}'$, σ varies in $\{1, 2, 3\}$, λ takes values $\{0.5, 0.2, -0.2, -0.5\}$, and n is chosen to be $\{50, 100, 200, 500\}$.[10] Several ways of generating W_n, (X_{n1}, X_{n2}), and u_n are considered. First, the values $\{x_{1i}\}$ or $\{x_{1,ir}\}$ of X_{n1}, and the values $\{x_{2i}\}$ or $\{x_{2,ir}\}$ of X_{n2} are,

`MRSAR-A`: $\{x_{1i}\} \overset{iid}{\sim} 10U(0,1)$, and $\{x_{2i}\} \overset{iid}{\sim} 5N(0,1)+5$, or
`MRSAR-B`: $\{x_{1,ir}\} = 5z_r + z_{ir}$, and $\{x_{2,ir}\} = v_r + v_{ir}$, or
`MRSAR-C`: $\{x_{1,ir}\} = (2z_r + z_{ir})/\sqrt{5}$, and $\{x_{2,ir}\} = (v_r + v_{ir})/\sqrt{2}$,

where in both `MRSAR-B` and `MRSAR-C`, $\{z_r, z_{ir}, v_r, v_{ir}\} \overset{iid}{\sim} N(0,1)$, across all i and r. Apparently, `MRSAR-A` gives iid X values, and `MRSAR-B` and `MRSAR-C` give non-iid X values, or different group means under group interaction, see Lee (2004a) and below for details.

Spatial layouts. Three general spatial layouts are considered in the Monte Carlo experiments. The first is based on Rook contiguity, the second is based on Queen contiguity and the third is based on the notion of group interactions. The methods used in generating these three spatial layouts are similar to those used in Yang (2010).

The detail for generating the W_n matrix under rook contiguity is as follows: (i) index the n spatial units by $\{1, 2, \cdots, n\}$, randomly permute these indices and then allocate them into a lattice of $k \times m(\geq n)$ squares, (ii) let $W_{ij} = 1$ if the index j is in a square which is on immediate left, or right, or above, or below the square which contains the index i, otherwise $W_{ij} = 0, i, j = 1, \cdots, n$, to form an $n \times n$ matrix, and (iii) divide each element of this matrix by its row sum to give W_n. Similarly, one generates the W_n matrix under Queen contiguity with additional neighbors sharing a common vertex with the unit of interest.

To generate the W_n matrix according to the group interaction scheme, suppose we have k groups of sizes $m_1, m_2, \cdots, m_k$. Define $W_n = \text{diag}\{W_j/(m_j - 1), j = 1, \cdots, k\}$, a matrix

[10] As in Lee (2007a), the maximization of the profile likelihood function is performed globally without imposing a restricted parameter space, such as λ lies in $(-1, 1)$, in our studies. This is important when the true λ value is negative and big, because QMLE is downward biased and a restricted lower bound, -0.9999 say, would result in the searching process to hit the lower bound quite often. This would in turn gives a wrong impression that when λ is negative, the QMLE becomes upward biased. It is believed that this is the reason for the incoherent Monte Carlo results of Bao and Ullah (2007a).

formed by placing the submatrices W_j along the diagonal direction, where W_j is an $m_j \times m_j$ matrix with ones on the off-diagonal positions and zeros on the diagonal positions. The group sizes $\{m_j\}$ can be the same or different, and independent or dependent on n, allowing for a full range of spatial scenarios considered in Lee (2004a). The details are as follows: (i) calculate the number of groups according to $k = K(n)$, and the approximate average group size $m = n/k$, (ii) generate the group sizes $(m_1, m_2, \cdots, m_k)$ according to a discrete distribution centered at m, and (iii) adjust the group sizes so that $\sum_{j=1}^{k} m_j = n$.[11]

In our Monte Carlo experiments, we use $K(n) = \text{Round}(n^{\epsilon})$ with $\epsilon = 0.35, 0.50$, and 0.75, representing respectively the situations where (a) there are few groups of many spatial units in each, (b) the number of groups and the sizes of the groups are of the same magnitude, and (c) there are many groups of few elements in each. Clearly, $h_n = O(n^{1-\epsilon})$. The group sizes are drawn from a discrete uniform distribution from $0.5m$ to $1.5m$.

Error distributions. To generate $u_n = \sigma e_n$, three distributions are considered: `dgp1`: the elements $\{e_i\}$ of e_n are iid standard normal, `dgp2`: $\{e_i\}$ are iid standardized normal mixture, and `dgp3`: $\{e_i\}$ are iid standardized log-normal. Specifically, for `dgp2`,

$$e_i = ((1 - \xi_i)Z_i + \xi_i \tau Z_i)/(1 - p + p * \sigma^2)^{0.5},$$

where $\xi \sim \text{Bernoulli}(p)$, and $Z_i \sim N(0, 1)$ independent of ξ. The parameter p represents the proportion of mixing the two normal populations. In our experiments, we choose $p = 0.1$, meaning that 90% of the random variates are from standard normal and the remaining 10% are from another normal population with standard deviation τ. We choose $\tau = 4$ to simulate the situation where there are gross errors in the data. For `dgp3`,

$$e_i = [\exp(Z_i) - \exp(0.5)]/[\exp(2) - \exp(1)]^{0.5},$$

which gives an error distribution that is both skewed and leptokurtic. The normal mixture gives an error distribution that is still symmetric like normal but leptokurtic. Other non-normal distributions, such as normal-gamma mixture and chi-square, are also considered and the results (available from the author upon request) exhibit a similar pattern.

Finite sample performance of the spatial estimators. We report the Monte Carlo mean, rmse and sd of $\hat{\lambda}_n$, $\hat{\lambda}_n^{\text{ba2}}$, $\hat{\lambda}_n^{\text{bc2}}$ and $\hat{\lambda}_n^{\text{bc3}}$ under various combinations of the values for

[11] Clearly, this design covers the scenario considered in Case (1991). Typical forms of $K(n)$ include $K(n) = n/m$ where m is a prespecified constant independent of n and $K(n) = \text{Round}(n^{\epsilon})$. Lee (2007b) shows that the group size variation plays an important role in the identification and estimation of econometric models with group interactions, contextual factors and fixed effects. Yang (2010) shows that it also plays an important role in the robustness of the LM test of spatial error components.

(n, λ, σ), the error distributions, and the spatial layouts. Each set of results is based on 10,000 Monte Carlo samples. Partial results are summarized in Tables 1-5, where Table 1 corresponds to the Queen contiguity spatial layout with the regressor values generated according to `MRSAR-A`, and Tables 2-5 correspond to the group interaction spatial layout with number of groups $k = \text{Round}(n^{0.5})$ or $\text{Round}(n^{0.35})$ and the values of the regressors are generated according to `MRSAR-B`. Some general observations are in order:

(i) the bias-corrected QMLEs $\hat{\lambda}_n^{\text{bc2}}$ and $\hat{\lambda}_n^{\text{bc3}}$ are in general nearly unbiased and clearly outperform the original QMLE $\hat{\lambda}_n$ and the bias-adjusted QMLE $\hat{\lambda}_n^{\text{ac}}$;

(ii) $\hat{\lambda}_n$ is always downward biased and some times the biasedness can be quite serious depending on the spatial layout and the error standard deviation, and $\hat{\lambda}_n^{\text{ac2}}$ does not seem to offer an improvement due to excluding the first term in (19);

(iii) $\hat{\lambda}_n^{\text{bc3}}$ improves over $\hat{\lambda}_n^{\text{bc2}}$ but is often marginal, suggesting that the second-order bias-correction is often sufficient for the SAR model;[12]

(iv) spatial layouts can have a huge impact on the finite sample performance of $\hat{\lambda}_n$ and $\hat{\lambda}_n^{\text{ac2}}$ - the stronger the spatial dependence the worse they perform;

(v) the error standard deviation σ also has a big impact on the finite sample performance of $\hat{\lambda}_n$ and $\hat{\lambda}_n^{\text{ac2}}$ - the bigger the σ is, the bigger are the biases, rmses and sds;

(vi) the value of λ also affects the finite sample performance of $\hat{\lambda}_n$ and $\hat{\lambda}_n^{\text{ac}}$ - generally speaking the more negative the λ is the larger are the biases, rmse and se; and finally

(vii) the sds of $\hat{\lambda}_n$, $\hat{\lambda}_n^{\text{bc2}}$ and $\hat{\lambda}_n^{\text{bc3}}$ agree quite well in general, which is consistent with the general reasoning given in Section 3.4.

(viii) the proposed bias-correction procedure works equally well for models with normal or nonnormal errors.

Bias of other parameter estimators. The biasedness of $\hat{\beta}_n \equiv \hat{\beta}_n(\hat{\lambda}_n)$, $\hat{\beta}_n^{\text{bc}} \equiv \hat{\beta}_n(\hat{\lambda}_n^{\text{bc2}})$, $\tilde{\sigma}_n^2 \equiv \frac{n}{n-k}\hat{\sigma}_n^2(\hat{\lambda}_n)$, and $\tilde{\sigma}_n^{2bc} \equiv \frac{n}{n-k}\hat{\sigma}_n^2(\hat{\lambda}_n^{\text{bc2}})$ are investigated as well. Partial results are summarized in Tables 6-7 under `MRSAR-A` and Tables 8-9 under `MRSAR-B`, all under group interaction spatial layouts. From the results, we see that $\hat{\beta}_n$ can be quite biased but $\hat{\beta}_n^{\text{bc}}$ is nearly unbiased, consistent with our theoretical arguments given in Section 3.3 that once $\hat{\lambda}_n$ is bias-corrected, the corresponding QMLE of β becomes automatically unbiased. Furthermore, Tables 6-7 show that the bias of $\hat{\beta}_n$ occurs only on the intercept, whereas Tables 8-9 show that the bias occurs on all elements of $\hat{\beta}_n$. In the former situation, the

[12] We note when $\lambda = -0.5$, $\sigma = 3$ and $n = 50$ (Tables 3 and 5), $\hat{\lambda}_n^{\text{bc3}}$ does not seem to improve over $\hat{\lambda}_n^{\text{bc2}}$, but this may well be due to the large variability of $\hat{\lambda}_n$ under these circumstances.

values of X_1 and X_2 are iid and hence their group means are the same, whereas in the latter situation they are not and as a result the group means of the values of X_1 and X_2 are different. It is interesting to note that both $\tilde{\sigma}_n^2$ and $\tilde{\sigma}_n^{2bc}$ are nearly unbiased. Again, the error distribution does not affect much the performance of the bias-corrected estimators.

The performance of the usual t-ratios. The behavior of the t-ratios for λ: $t(\lambda) = (\hat{\lambda}_n - \lambda)/(\widehat{\text{AVar}}(\hat{\lambda}_n))^{\frac{1}{2}}$ and $t^{\text{bc}}(\lambda) = (\hat{\lambda}_n^{\text{bc2}} - \lambda)/(\widehat{\text{V}}_2(\hat{\lambda}_n))^{\frac{1}{2}}$, and the t-ratios for β_1: $t(\beta_1) = (\hat{\beta}_{n1} - \beta_1)/(\widehat{\text{AVar}}(\hat{\beta}_{n1}))^{\frac{1}{2}}$ and $t^{\text{bc}}(\beta_1) = (\hat{\beta}_{n1}^{bc} - \beta_1)/(\widehat{\text{AVar}}(\hat{\beta}_{n1}))^{\frac{1}{2}}$ is investigated, where $\widehat{\text{AVar}}$ denotes the asymptotic variance, and $\widehat{\text{V}}_2(\hat{\lambda}_n)$ is the bootstrap estimate of the second-order variance given in (25). Partial Monte Carlo results are summarized in Tables 10 and 11. From the results we see that, in terms of mean, standard deviation, and the nominal 5% tail probability, the t-ratio for λ with the second-order corrections on both the estimator and its variance improves substantially over the usual t-ratio using the original QMLE and asymptotic variance, when referred to the limiting distribution $N(0, 1)$. The t-ratios for the β coefficients also show some improvements in making inference for β. Once again, the exact error distribution does not may a significant difference on the performance of the t-ratios.

We have tried using (23) instead of (25) for obtaining the second-order se of $\hat{\lambda}$, but found that it can give rise to negative estimates of the variance when spatial dependence is strong, error standard deviation is big, and sample size is small. This is because when the variabilities of the random quantities are large, the second-order term $2\text{Cov}(a_{-1/2,a_{-1}})$ in (23) can be larger in magnitude than the first order term $\text{Var}(a_{-1/2})$ in a certain bootstrap sample resulting a negative variance estimate. However, with the addition of a third-order term $\text{Var}(a_{-1})$ as in (25), the variance estimate is guaranteed to be positive.

4.2 Monte Carlo Experiment II

The Monte Carlo experiment conducted by Bao and Ullah (2007a) is replicated and extended for two purposes: (i) to compare our bootstrap-based bias correction with their analytical bias correction, and (ii) to investigate the cause of incoherent Monte Carlo results of Bao and Ullah (2007a) for the case of $\lambda = -0.9$ and $J = 10$, where J is the number of neighbors each spatial unit has.

The empirical means and rmses for the four estimators: the MLE $\hat{\lambda}_n$, the bias-adjusted MLE $\hat{\lambda}_n^{\text{ba2}}$ given in (22), the bias-corrected MLE $\hat{\lambda}_n^{\text{bc2}}$ given in (21), and the analytically bias-corrected MLE $\hat{\lambda}_n^{\text{bu}}$ of Bao and Ullah (2007a), are summarized in Table 12. The results show that (i) $\hat{\lambda}_n^{\text{ba2}}$ and $\hat{\lambda}_n^{\text{bc2}}$ are the same as predicted by the theory and they behave very similarly to $\hat{\lambda}_n^{\text{bu}}$ which is based on the analytical formulas of Bao and Ullah (2007a), showing the bootstrap procedure works well, and (ii) the results show a very coherent behavior of the MLE and the bias-corrected MLEs in that the MLE is almost unbiased when J is small and is downward biased when J is big. In cases where the MLE is biased, the bias-corrected estimators always be able to correct the bias. However, the Monte Carlo results of Bao and Ullah show a different picture for the cases where J is large and λ is large and negative: the MLE is upward biased and the bias-correction does not work. A possible explanation for this can be found in Footnote 10.

4.3 Monte Carlo Experiment III

The Monte Carlo experiments of Lee (2004) are rerun and extended again for two purposes: i) to further compare MLEs (estimators under normal errors) and second-order bias-corrected MLEs under a model without intercept, and ii) to amend the anomalies in his Monte Carlo results. Under the identical set-up but with more Monte Carlo replications (10,000 vs 400) for each case, we found that with the number of districts $R = (30, 60, 120)$ and the number of members in each district $m = (2, 3, 5, 1020, 50, 100)$, the MLEs all perform very well in all cases, showing a big contrast to the results of Lee (2004a). However, when the values for R and m are switched, the results obtained are more in line with those of Lee (2004a) in terms of bias but not in terms of sd.

Table 13 corresponds to Tables I & II in Lee (2004a). Tables 14 summarize the results of the Monte Carlo experiments when the values for R and m are switched, i.e., $R = (3, 5, 10, 20, 50, 100)$ and $m = (30, 60, 120)$. The results in Tables 14 are now more in line with Lee's results in terms of bias but not in terms of sd of $\hat{\lambda}_n$. The results further show that for a given R value, increasing the m value does not necessarily reduce the bias of $\hat{\lambda}_n$. However, for a given m, increasing R clearly improves the performance of the estimators with both bias and sd significantly reduced.

From the theoretical point of view, our results make more sense as increasing m enlarges the degree of spatial dependence, making the estimation of the spatial parameter harder (or at least not easier). In contrast, increasing R clearly reduces the 'degree' of spatial dependence, making the estimation of the spatial parameter much easier.

5 Conclusions

To address the biasedness issue in a model containing nonlinear, location as well as scale parameters, one needs only to focus on the estimation of the nonlinear parameter and use the concentrated estimating equation to obtain higher-order expansions to achieve bias-correction. This turns a multidimensional problem to a single dimension and greatly simplifies the higher-order expansions. It is argued that for these abstract formulas to be practically useful, it is necessary to have a feasible method for estimating the various expectations in the formulas. Thus, a simple bootstrap procedure is introduced. These ideas and methods are explored in a full detail in the context of a spatial autoregressive model. Monte Carlo results show that this approach is quite effective in that it almost eliminates the bias of the QMLE, which can be quite large when spatial dependence is strong.

The methods introduced in this paper can be applied to other models of similar nature. For example, a panel model with a spatial lag where the QMLE of the spatial lag parameter would incur biasness, linear regression with a response transformation where the QMLE of the transformation parameter may incur biasness, a dynamic regression model where the QMLE of the dynamic parameter may be biased, etc. While applying the proposed methods to other models of a similar structure are interesting topics for future research, a more detailed study on inferences following a bias-corrected estimation of the spatial parameter may be of a immediate interest. These inferences include tests and confidence intervals for spatial effects, tests and confidence intervals for regression effects, predicting future response values, etc., based on higher-order corrections on the mean as well as on the standard errors.

Our methods can in principle be generalized to allow for asymptotic (first-order) bias based on a CEE, and/or to the cases where there are several (though still only a few) bias-inducing parameters in the model. Typical models of both features are the panel models (dynamic or nonlinear) with fixed effects, and in these cases, it would be interesting to extend our methods to give higher-order bias reduction to the problems considered in Hahn and Kuersteiner (2002) and to offer an alternative to the jackknife and analytical bias reduction method of Hahn and Newey (2004) which is based on an iid data set-up. Other models of more than one bias-inducing parameters include the SAR model where the disturbances also follow a SAR process giving two spatial parameters, spatial panel models with random effects, etc. We are planning to pursue these issues in future research.

Appendix A: Proofs of the Results in Section 2

Proof of Theorem 2.1: Assumption B allows the Taylor expansions of $\tilde{\psi}_n(\hat{\lambda}_n) = 0$ around λ_0 to be carried out up to third-order, and Assumptions A and D guarantee that the errors in the Taylor approximations are of order $O_p(n^{-1})$, $O_p(n^{-3/2})$, and $O_p(n^{-2})$, respectively, for the 1st-, 2nd- and 3rd-order Taylor expansion. We thus have,

$$\begin{aligned}
0 &= \tilde{\psi}_n + H_{1n}(\hat{\lambda}_n - \lambda_0) + O_p(n^{-1}), \\
0 &= \tilde{\psi}_n + H_{1n}(\hat{\lambda}_n - \lambda_0) + \frac{1}{2}H_{2n}(\hat{\lambda}_n - \lambda_0)^2 + O_p(n^{-3/2}), \\
0 &= \tilde{\psi}_n + H_{1n}(\hat{\lambda}_n - \lambda_0) + \frac{1}{2}H_{2n}(\hat{\lambda}_n - \lambda_0)^2 + \frac{1}{6}H_{2n}(\hat{\lambda}_n - \lambda_0)^3 + O_p(n^{-2}),
\end{aligned}$$

which give, as $-H_{1n}^{-1} = O_p(1)$ from Assumption C,

$$\hat{\lambda}_n - \lambda_0 = -H_{1n}^{-1}\tilde{\psi}_n + O_p(n^{-1}), \tag{A-1}$$

$$\hat{\lambda}_n - \lambda_0 = -H_{1n}^{-1}\tilde{\psi}_n - \frac{1}{2}H_{1n}^{-1}H_{2n}(\hat{\lambda}_n - \lambda_0)^2 + O_p(n^{-3/2}), \tag{A-2}$$

$$\hat{\lambda}_n - \lambda_0 = -H_{1n}^{-1}\tilde{\psi}_n - \frac{1}{2}H_{1n}^{-1}H_{2n}(\hat{\lambda}_n - \lambda_0)^2 - \frac{1}{6}H_{1n}^{-1}H_{2n}(\hat{\lambda}_n - \lambda_0)^3 + O_p(n^{-2}). \tag{A-3}$$

Under Assumptions B and C, $\Omega_n = -\mathrm{E}(H_{1n})^{-1} = O(1), H_{1n}^{-1} = O_p(1)$, and $H_{1n}^{\circ} = H_{1n} - \mathrm{E}(H_{1n}) = O_p(n^{1/2})$. These conditions lead to the following result

$$-H_{1n}^{-1} = (\Omega_n^{-1} - H_{1n}^{\circ})^{-1} = (1 - \Omega_n H_{1n}^{\circ})^{-1}\Omega_n = \Omega_n + \Omega_n^2 H_{1n}^{\circ} + \Omega_n^3 H_{1n}^{\circ 2} + O_p(n^{-3/2}),$$

which reduces to $-H_{1n}^{-1} = \Omega_n + \Omega_n^2 H_{1n}^{\circ} + O_p(n^{-1})$, or $= \Omega_n + O_p(n^{-1/2})$. Substituting $-H_{1n}^{-1} = \Omega_n + O_p(n^{-1/2})$ into (A-1) gives a first-order stochastic expansion for $\hat{\lambda}_n$,

$$\hat{\lambda}_n - \lambda_0 = \Omega_n\tilde{\psi}_n + O_p(n^{-1}) = a_{-1/2} + O_p(n^{-1}). \tag{A-4}$$

Substituting (A-4) into (A-2) for $\hat{\lambda}_n - \lambda_0$, $\Omega_n + \Omega_n^2 H_{1n}^{\circ} + O_p(n^{-1})$ into (A-2) for the first $-H_{1n}^{-1}$, and $\Omega_n + O_p(n^{-\frac{1}{2}})$ into (A-2) for the second $-H_{1n}^{-1}$, give a second-order stochastic expansion for $\hat{\lambda}_n$ which takes the form after some algebra,

$$\hat{\lambda}_n - \lambda_0 = a_{-1/2} + a_{-1} + O_p(n^{-3/2}), \tag{A-5}$$

where $a_{-1} = \Omega_n H_{1n}^{\circ} a_{-1/2} + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2}^2)$. Finally, substituting (A-5) into (A-3) for $\hat{\lambda}_n - \lambda_0$ in the second term, (A-4) into (A-3) for $\hat{\lambda}_n - \lambda_0$ in the third term, $\Omega_n + \Omega_n^2 H_{1n}^{\circ} + \Omega_n^3 H_{1n}^{\circ 2} + O_p(n^{-3/2})$ into (A-3) for the first $-H_{1n}^{-1}$, $\Omega_n + \Omega_n^2 H_{1n}^{\circ} + O_p(n^{-1})$ into (A-3) for the second $-H_{1n}^{-1}$, and $\Omega_n + O_p(n^{-1/2})$ into (A-3) for the third $-H_{1n}^{-1}$, give a third-order stochastic expansion for $\hat{\lambda}_n$,

$$\hat{\lambda}_n - \lambda_0 = a_{-1/2} + a_{-1} + a_{-3/2} + O_p(n^{-2}), \tag{A-6}$$

where $a_{-3/2} = \Omega_n H^{\circ}_{1n} a_{-1} + \frac{1}{2}\Omega_n H^{\circ}_{2n}(a^2_{-1/2}) + \Omega_n \mathrm{E}(H_{2n})(a_{-1/2}a_{-1}) + \frac{1}{6}\Omega_n \mathrm{E}(H_{3n})(a^3_{-1/2})$.

Proof of Corollary 2.1: We have $b_{-1} = \mathrm{E}(a_{-1/2}) + \mathrm{E}(a_{-1})$, and $b_{-3/2} = \mathrm{E}(a_{-3/2})$. The result follows if the expected error term $E(O_p(n^{-2})) = O(n^{-2})$, which follows the assumption in the corollary.

Proof of Corollary 2.2: We have $\mathrm{MSE}(\hat{\lambda}_n) = E[(a_{-1/2} + a_{-1} + a_{-3/2} + O_p(n^{-2})^2]$. Under the assumption that a quantity bounded in probability has a finite expectation stated in the corollary, we have $\mathrm{MSE}(\hat{\lambda}_n) = m_{-1} + m_{-3/2} + m_{-2} + O(n^{-5/2})$, where $m_{-1} = \mathrm{E}(a^2_{-1/2})$, $m_{-3/2} = 2\mathrm{E}(a_{-1/2}a_{-1})$, and $m_{-2} = 2\mathrm{E}[(a_{-1/2}a_{-3/2}) + a^2_{-1}]$. The rest is straightforward.

Proof of Corollary 2.3: Straightforward from the proofs of Corollaries 2.1 and 2.2.

Appendix B: Proofs of the Results in Section 3

To prove the results of Section 3, we need the following lemmas.

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

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

(ii) the elements of A_n are uniformly bounded and $tr(A_n) = O(n)$, and

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

Lemma B.2 (Lee, 2004a, p.1918): *Let X_n be an $n \times k$ matrix such that (i) its elements are uniformly bounded; and (ii) $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ exists and is nonsingular. Then the projectors $P_n = X_n(X_n'X_n)^{-1}X_n'$ and $M_n = I_n - X_n(X_n'X_n)^{-1}X_n'$ are uniformly bounded in both row and column sums.*

Lemma B.3 (Lemma A.9, Lee, 2004b): *Let $\{A_n\}$ be a sequence of $n \times n$ matrices that are uniformly bounded in both row and column sums. For M_n defined in Lemma B.2,*

(i) $\mathrm{tr}(M_n A_n) = \mathrm{tr}(A_n) + O(1)$

(ii) $\mathrm{tr}(A_n' M_n A_n) = \mathrm{tr}(A_n' A_n) + O(1)$

(iii) $\mathrm{tr}[(M_n A_n)^2] = \mathrm{tr}(A_n^2) + O(1)$, *and*

(iv) $\mathrm{tr}[(A_n' M_n A_n)^2] = \mathrm{tr}[(M_n A_n' A_n)^2] = \mathrm{tr}[A_n' A_n)^2] + O(1)$

Furthermore, if the elements $a_{n,ij}$ of A_n are $O(h_n^{-1})$ uniformly in all i and j, then, *(v)* $\mathrm{tr}^2(M_nA_n) = \mathrm{tr}^2(A_n) + O(\frac{n}{h_n})$ *and*

(vi) $\mathrm{tr}\sum_{i=1}^n((M_nA_n)_{ii})^2 = \sum_{i=1}^n((M_nA_n)_{ii})^2 + O(h_n^{-1})$,

where $(M_nA_n)_{ii}$ *is the ith diagonal element of* M_nA_n.

Lemma B.4 (Lemma A.12, Lee, 2004b, extended): *Let* $\{A_n\}$ *be a sequence of* $n \times n$ *matrices that are uniformly bounded in either row or column sums. Suppose that the elements* $a_{n,ij}$ *of* A_n *are* $O(h_n^{-1})$ *uniformly in all i and j. Let* u_n *be a random n-vector of iid elements with finite 4th moment, and* b_n *be a constant n-vector of which the elements are of uniform order* $O(h_n^{-1/2})$. *Then (i)* $\mathrm{E}(u_n'A_nu_n) = O(\frac{n}{h_n})$, *(ii)* $\mathrm{Var}(u_n'A_nu_n) = O(\frac{n}{h_n})$, *(iii)* $u_n'A_nu_n = O_p(\frac{n}{h_n})$, *(iv)* $u_n'A_nu_n - \mathrm{E}(u_n'A_nu_n) = O_p((\frac{n}{h_n})^{\frac{1}{2}})$, *and (v)* $u_n'A_nb_n = O_p((\frac{n}{h_n})^{\frac{1}{2}})$.

Note that the results (iv) and (v) in Lemma B.4 extend Lemma A.12 of Lee (2004b), where (iv) follows directly from the generalized Chebyshev inequality and the result (ii): $P(\sqrt{\frac{h_n}{n}}|u_n'A_nu_n - \mathrm{E}(u_n'A_nu_n)| \geq M) \leq \frac{1}{M^2}\frac{h_n}{n}\mathrm{Var}(u_n'A_nu_n) = \frac{1}{M^2}O(1)$; and (v) follows from the generalized Chebyshev inequality: $P(\sqrt{\frac{h_n}{n}}|u_n'A_nb_n)| \geq M) \leq \frac{1}{M^2}\frac{h_n}{n}\mathrm{Var}(u_n'A_nb_n) = \frac{1}{M^2}\frac{h_n}{n}b_n'A_n'A_nb_n = \frac{1}{M^2}O(1)$.

Lemma B.5 (Kelejian and Prucha, 2001, p.227, extended): *Let* $\{A_n\}$ *be an* $n \times n$ *matrix of elements* $\{a_{n,ij}\}$, b_n *be an* $n \times 1$ *vector of elements* $\{b_{n,i}\}$, *and* u_n *an* $n \times 1$ *random vector of iid elements, having mean zero, variance* σ_0^2, *skewness* γ, *and excess kurtosis* κ. *Let* $Q_n = u_n'A_nu_n + b_n'u_n$. *Then,*

(i) $E(Q_n) = \sigma_0^2\mathrm{tr}(A_n)$,

(ii) $Var(Q_n) = \sigma_0^4\mathrm{tr}(A_nA_n' + A_n^2) + \sigma_0^4\kappa\sum_{i=1}^n a_{n,ii}^2 + \sigma_0^2\sum_{i=1}^n b_{n,i}^2 + 2\sigma_0^3\gamma\sum_{i=1}^n b_{n,i}a_{n,ii}$.

Furthermore, if $\{a_{n,ij}\}$ *are of uniform order* $O_p(h_n^{-1})$, $\{b_{n,i}\}$ *are of uniform order* $O_p(h_n^{-\frac{1}{2}})$, *and* $\{A_n\}$ *are uniformly bounded in either row or column sums, then*

(iii) $E(Q_n) = O(\frac{n}{h_n})$, *and*

(iv) $Var(Q_n) = O(\frac{n}{h_n})$.

Proof of Lemma 3.1. By a Taylor series expansion (of $\hat{\sigma}_{n0}^{-2}$ around σ_0^2), we have,

$$\begin{aligned} h_nR_{1n} &= \tfrac{h_n}{n}\hat{\sigma}_{n0}^{-2}Y_n'A_n'M_nW_nY_n \\ &= \tfrac{h_n}{n}\sigma_0^{-2}Y_n'A_n'M_nW_nY_n - \tfrac{h_n}{n}\sigma_0^{-4}(Y_n'A_n'M_nW_nY_n)(\hat{\sigma}_{n0}^2 - \sigma_0^2) \\ &\quad -\tfrac{h_n}{n}(\bar{\sigma}_{n0}^{-4} - \sigma_0^{-4})(Y_n'A_n'M_nW_nY_n)(\hat{\sigma}_{n0}^2 - \sigma_0^2), \end{aligned}$$

where $\bar{\sigma}_{n0}^2$ lies between $\hat{\sigma}_{n0}^2$ and σ_0^2. Now, $Y_n'A_n'M_nW_nY_n = u_n'M_nG_nu_n + u_n'M_n\eta_n$. Lemma B.4(iii) implies $\frac{h_n}{n}u_n'M_nG_nu_n = O_p(1)$ and Lemma B.4(v) implies $\frac{h_n}{n}u_n'M_n\eta_n = O_p((\frac{h_n}{n})^{\frac{1}{2}})$. It follows that $\frac{h_n}{n}Y_n'A_n'M_nW_nY_n = \frac{h_n}{n}u_n'M_nG_nu_n + O_p((\frac{h_n}{n})^{\frac{1}{2}}) = O_p(1)$. Furthermore, Assumptions 1-6 lead to $\hat{\sigma}_{n0}^2 - \sigma_0^2 = O_p(n^{-\frac{1}{2}})$, which implies $\bar{\sigma}_{n0}^{-4} - \sigma_0^{-4} = O_p(n^{-\frac{1}{2}})$. Thus,

$$h_n R_{1n} = \frac{h_n}{\sigma_0^2 n} u_n' M_n G_n u_n + O_p((h_n/n)^{\frac{1}{2}}) \tag{B-1}$$

On the other hand,

$$\begin{aligned} \mathrm{E}(h_n R_{1n}) &= \tfrac{h_n}{n}\sigma_0^{-2}\mathrm{E}(Y_n' A_n' M_n W_n Y_n) - \tfrac{h_n}{n}\sigma_0^{-4}\mathrm{E}[(Y_n' A_n' M_n W_n Y_n)(\hat{\sigma}_{n0}^2 - \sigma_0^2)] \\ &\quad - \tfrac{h_n}{n}\mathrm{E}[(\bar{\sigma}_{n0}^{-4} - \sigma_0^{-4})(Y_n' A_n' M_n W_n Y_n)(\hat{\sigma}_{n0}^2 - \sigma_0^2)], \end{aligned}$$

where the first term becomes $\frac{h_n}{n}\mathrm{tr}(M_n G_n)$, the last term is $O(n^{-1})$ under Assumption 7. Now, by Cauchy-Schwarz inequality, and using Lemma B.5, we have

$$\begin{aligned} |\mathrm{E}[(Y_n' A_n' M_n W_n Y_n)(\hat{\sigma}_{n0}^2 - \sigma_0^2)]| &\leq \tfrac{1}{n}\{\mathrm{Var}(Y_n' A_n' M_n W_n Y_n)\mathrm{Var}(u_n' M_n u_n - n\sigma_0^2)\}^{\frac{1}{2}} \\ &= \tfrac{1}{n}\{\mathrm{Var}(u_n' M_n G_n u_n + u_n' M_n \eta_n)\mathrm{Var}(u_n' M_n u_n - n\sigma_0^2)\}^{\frac{1}{2}} \\ &= \tfrac{1}{n}\{O(\tfrac{n}{h_n})O(n)\}^{\frac{1}{2}} = O(h_n^{-\frac{1}{2}}). \end{aligned}$$

These lead to

$$\mathrm{E}(h_n R_{1n}) = \frac{h_n}{n}\mathrm{tr}(M_n G_n) + O(h_n^{1/2}/n). \tag{B-2}$$

Taking difference between (B-1) and (B-2) and using Lemma B.4(iv), we obtain $h_n R_{1n} - \mathrm{E}(h_n R_{1n}) = \frac{h_n}{\sigma_0^2 n} u_n' M_n G_n u_n - \frac{h_n}{n}\mathrm{tr}(M_n G_n) + O_p((\frac{h_n}{n})^{\frac{1}{2}}) - O(\frac{h_n^{1/2}}{n}) = O_p((\frac{h_n}{n})^{\frac{1}{2}})$.

Following similar arguments, we show that

$$\begin{aligned} h_n R_{2n} &= \frac{h_n}{\sigma_0^2 n}(u_n' G_n' M_n G_n u_n + \eta_n' M_n \eta_n) + O_p((h_n/n)^{\frac{1}{2}}), & \text{(B-3)} \\ \mathrm{E}(h_n R_{2n}) &= \frac{h_n}{n}\mathrm{tr}(G_n' M_n G_n) + \frac{h_n}{\sigma_0^2 n}\eta_n' M_n \eta_n + O(h_n^{1/2}/n), & \text{(B-4)} \end{aligned}$$

which gives the result in Lemma 3.1(ii) upon applying Lemma B.4(iv). *Q.E.D.*

Proof of Theorem 3.1. Clearly, the $\tilde{\psi}(\lambda)$ function given in (12) is differentiable for λ in a neighborhood of λ_0 with its first three derivatives $H_{rn}(\lambda), r = 1, 2$, and 3, given in (13)-(15). These allow us to implement the following third-order Taylor expansion:

$$\begin{aligned} 0 = \tilde{\psi}_n(\hat{\lambda}_n) &= \tilde{\psi}_n + H_{1n}(\hat{\lambda}_n - \lambda_0) + \frac{1}{2}H_{2n}(\hat{\lambda}_n - \lambda_0)^2 + \frac{1}{6}H_{3n}(\hat{\lambda}_n - \lambda_0)^3 \\ &\quad + \frac{1}{6}[H_{3n}(\bar{\lambda}) - H_{3n}](\hat{\lambda}_n - \lambda_0)^3, \end{aligned}$$

where $\bar{\lambda}$ lies between $\hat{\lambda}_n$ and λ_0. Under Assumptions 1-6, $\hat{\lambda}_n$ is $\sqrt{n/h_n}$-consistent. Incorporating h_n and following the arguments leading to the result of Theorem 2.1, the result of Theorem 3.1 follows if the following results hold:

(a) $\tilde{\psi}_n = O((\frac{h_n}{n})^{\frac{1}{2}})$ and $\mathrm{E}(\tilde{\psi}_n) = O(\frac{h_n}{n})$;

(b) $\mathrm{E}(H_{rn}) = O(1)$ and $H^\circ_{rn} = O_p((\frac{h_n}{n})^{\frac{1}{2}}), r = 1, 2, 3$;
(c) $H^{-1}_{1n} = O_p(1)$ and $\mathrm{E}(H_{1n})^{-1} = O(1)$; and
(d) $H_{3n}(\bar{\lambda}) - H_{3n} = O_p((\frac{h_n}{n})^{\frac{1}{2}})$.

First, Lemma B.1 and Assumptions 3 and 4 give $h_n T_{rn} = O(1), r = 1, 2, 3$.

For (a), by (B-1), $\tilde{\psi}_n = -h_n T_{0n} + h_n R_{1n} = -h_n T_{0n} + \frac{h_n}{\sigma_0^2 n} u'_n M_n G_n u_n + O_p((\frac{h_n}{n})^{\frac{1}{2}})$; by Lemma B.3(i), $\mathrm{tr}(M_n G_n) = \mathrm{tr}(G_n) + O(1) = nT_{0n} + O(1)$. Thus, $\tilde{\psi}_n = O_p((\frac{h_n}{n})^{\frac{1}{2}})$. By (B-2), $\mathrm{E}(\tilde{\psi}_n) = -h_n T_{0n} + \frac{h_n}{n}\mathrm{tr}(M_n G_n) + O(\frac{h_n^{1/2}}{n})$. By Lemma B.3(i), $\mathrm{tr}(M_n G_n) = \mathrm{tr}(G_n) + O(1)$. It follows that $\mathrm{E}(\tilde{\psi}_n) = O(\frac{h_n}{n})$.

For (b), we have under Assumption 8,

$$
\begin{aligned}
O((\tfrac{h_n}{n})^{\frac{1}{2}}) &= h_n^2 \mathrm{E}(R_{1n} - \mathrm{E}R_{1n})^2 = h_n^2 \mathrm{E}(R_{1n}^2) - h_n^2 (\mathrm{E}R_{1n})^2 \\
&\Rightarrow \mathrm{E}((h_n R_{1n})^2) = \mathrm{E}^2(h_n R_{1n}) + O((\tfrac{h_n}{n})^{\frac{1}{2}}). \\
O((\tfrac{h_n}{n})^{\frac{1}{2}}) &= h_n^2 \mathrm{E}(R_{2n} - \mathrm{E}R_{2n})^2 = h_n^2 \mathrm{E}(R_{2n}^2) - h_n^2 (\mathrm{E}R_{2n})^2 \\
&\Rightarrow \mathrm{E}((h_n R_{2n})^2) = (\mathrm{E}h_n R_{2n})^2 + O((\tfrac{h_n}{n})^{\frac{1}{2}}). \\
O((\tfrac{h_n}{n})^{\frac{1}{2}}) &= h_n^3 \mathrm{E}(R_{1n} - \mathrm{E}R_{1n})^3 \\
&= h_n^3 \mathrm{E}(R_{1n}^3 - 3R_{1n}^2 \mathrm{E}(R_{1n}) + 3R_{1n}(\mathrm{E}R_{1n})^2 - (\mathrm{E}R_{1n})^3) \\
&= h_n^3 \mathrm{E}(R_{1n}^3) - h_n^3 (\mathrm{E}R_{1n})^3 + O((\tfrac{h_n}{n})^{\frac{1}{2}}) \\
&\Rightarrow \mathrm{E}((h_n R_{1n})^3) = (\mathrm{E}h_n R_{1n})^3 + O((\tfrac{h_n}{n})^{\frac{1}{2}}).
\end{aligned}
$$

Similarly, one shows

$$
\begin{aligned}
&\mathrm{E}((h_n R_{1n})^4) = (\mathrm{E}h_n R_{1n})^4 + O((\tfrac{h_n}{n})^{\frac{1}{2}}), \\
&\mathrm{E}((h_n R_{2n})^2) = (\mathrm{E}h_n R_{2n})^2 + O((\tfrac{h_n}{n})^{\frac{1}{2}}), \\
&\mathrm{E}((h_n R_{1n})^s (h_n R_{2n})) = (\mathrm{E}h_n R_{1n})^s (\mathrm{E}h_n R_{2n}) + O((\tfrac{h_n}{n})^{\frac{1}{2}}), s = 1, 2.
\end{aligned}
$$

Now, Lemma 3.1 implies

$$
\begin{aligned}
&(h_n R_{1n})^s = (\mathrm{E}h_n R_{1n})^s + O_p((\tfrac{h_n}{n})^{\frac{1}{2}}), s = 2, 3, 4, \\
&(h_n R_{2n})^2 = (\mathrm{E}h_n R_{2n})^2 + O_p((\tfrac{h_n}{n})^{\frac{1}{2}}), \\
&(h_n R_{1n})^s (h_n R_{2n}) = (\mathrm{E}h_n R_{1n})^s (\mathrm{E}h_n R_{2n}) + O_p((\tfrac{h_n}{n})^{\frac{1}{2}}), s = 1, 2.
\end{aligned}
$$

As $h_n T_{rn} = O(1), r = 1, 2, 3$, the above results lead immediately to

$$
\begin{aligned}
&\mathrm{E}(H_{rn}) = O(1), r = 1, 2, 3, \qquad H_{1n} - \mathrm{E}(H_{1n}) = O_p((\tfrac{h_n}{n})^{\frac{1}{2}}), \\
&H_{2n} - \mathrm{E}(H_{2n}) = O_p((\tfrac{1}{h_n n})^{\frac{1}{2}}), \quad H_{3n} - \mathrm{E}(H_{3n}) = O_p((\tfrac{1}{h_n n})^{\frac{1}{2}}).
\end{aligned}
$$

For (c), from (B-2), (B-4), and the result $\mathrm{E}((h_n R_{1n})^2) = \mathrm{E}^2(h_n R_{1n}) + O((\frac{h_n}{n})^{\frac{1}{2}})$ obtained above, we have

$$
\begin{aligned}
\mathrm{E}(H_{1n}) &= -h_nT_{1n} - \mathrm{E}(h_nR_{2n}) + \tfrac{2}{h_n}\mathrm{E}((h_nR_{1n})^2) \\
&= -h_nT_{1n} - \tfrac{h_n}{n}\mathrm{tr}(G_n'M_nG_n) - \tfrac{h_n}{\sigma_0^2 n}\eta_n'M_n\eta_n - O(\tfrac{h_n^{1/2}}{n}) \\
&\quad + \tfrac{2}{h_n}(\tfrac{h_n}{n}\mathrm{tr}(M_nG_n) + O(\tfrac{h_n^{1/2}}{n}))^2 \\
&= -h_nT_{1n} - \tfrac{h_n}{n}\mathrm{tr}(G_n'M_nG_n) - \tfrac{h_n}{\sigma_0^2 n}\eta_n'M_n\eta_n + \tfrac{2}{h_n}(\tfrac{h_n}{n}\mathrm{tr}(M_nG_n))^2 + O(\tfrac{h_n^{1/2}}{n}) \\
&= -\tfrac{h_n}{n}\mathrm{tr}(G_n^2) - \tfrac{h_n}{n}\mathrm{tr}(G_n'G_n) - \tfrac{h_n}{\sigma_0^2 n}\eta_n'M_n\eta_n + \tfrac{2}{h_n}(\tfrac{h_n}{n}\mathrm{tr}(G_n))^2 + O(\tfrac{h_n}{n}) \\
&= -\tfrac{h_n}{n}\mathrm{tr}((G_n - T_{0n}I_n)^2) - \tfrac{h_n}{n}\mathrm{tr}((G_n - T_{0n}I_n)'(G_nT_{0n}I_n)) - \tfrac{h_n}{\sigma_0^2 n}\eta_n'M_n\eta_n + O(\tfrac{h_n}{n}).
\end{aligned}
$$

This shows that $\mathrm{E}(H_{1n}) < 0$ for n sufficiently large and thus $\mathrm{E}(H_{1n})^{-1} = O(1)$. As $H_{1n} = \mathrm{E}(H_{1n}) + O_p((\frac{h_n}{n})^{\frac{1}{2}})$, we have $H_{1n}^{-1} = O_p(1)$.

For (d), we have

$$
\begin{aligned}
\hat{\sigma}_n^2(\bar{\lambda}) &= \tfrac{1}{n}Y_n'A_n'(\bar{\lambda})M_nA_n(\bar{\lambda})Y_n \\
&= \tfrac{1}{n}Y_n'A_n'M_nA_nY_n - 2(\bar{\lambda} - \lambda_0)\tfrac{1}{n}Y_n'A_n'M_nW_nY_n + (\bar{\lambda} - \lambda_0)^2\tfrac{1}{n}Y_n'W_n'M_nW_nY_n \\
&= \hat{\sigma}_{n0}^2 - 2(\bar{\lambda} - \lambda_0)O_p(h_n^{-1}) + (\bar{\lambda} - \lambda_0)^2O_p(h_n^{-1}) \\
&= \hat{\sigma}_{n0}^2 + O_p((h_nn)^{\frac{1}{2}}), \text{ and thus}
\end{aligned}
$$

$$
\begin{aligned}
h_nR_{1n}(\bar{\lambda}) &= \hat{\sigma}_n^{-2}(\bar{\lambda})\tfrac{h_n}{n}Y_n'A_n'(\bar{\lambda})M_nW_nY_n \\
&= \hat{\sigma}_n^{-2}(\bar{\lambda})\tfrac{h_n}{n}Y_n'A_n'M_nW_nY_n - \hat{\sigma}_n^{-2}(\bar{\lambda})(\bar{\lambda} - \lambda_0)\tfrac{h_n}{n}Y_n'W_n'M_nW_nY_n \\
&= (h_nR_{1n} + O_p((h_nn)^{\frac{1}{2}})) - O_p((\tfrac{h_n}{n})^{\frac{1}{2}}) \\
&= h_nR_{1n} + O_p((\tfrac{h_n}{n})^{\frac{1}{2}}).
\end{aligned}
$$

Similarly, one shows that $h_nR_{2n}(\bar{\lambda}) = h_nR_{2n} + O_p((\frac{h_n}{n})^{\frac{1}{2}})$. By the mean value theorem, $h_nT_{3n}(\bar{\lambda}) = \frac{h_n}{n}\mathrm{tr}(G_n^4(\bar{\lambda})) = \frac{h_n}{n}\mathrm{tr}(G_n^4) + 4\frac{h_n}{n}\mathrm{tr}(G_n^3(\check{\lambda}))(\bar{\lambda} - \lambda_0)$, where $\check{\lambda}$ lies between $\bar{\lambda}$ and λ_0. By Assumption 4 and Lemma B.1, $\frac{h_n}{n}\mathrm{tr}(G_n^3(\check{\lambda})) = O(1)$. It follows that $h_nT_{3n}(\bar{\lambda}) - h_nT_{3n} = O_p((\frac{h_n}{n})^{\frac{1}{2}})$. These lead to $H_{3n}(\bar{\lambda}) - H_{3n} = O_p((\frac{h_n}{n})^{\frac{1}{2}})$. *Q.E.D.*

Proof of Corollary 3.1. Straightforward.

Proof of Corollary 3.2. Straightforward.

Proof of Corollary 3.3. Straightforward.

References

[1] Amemiya, T. (1985). *Advanced Econometrics.* Cambridge, Massachusetts: Harvard University Press.

[2] Anselin L. (1988). *Spatial Econometrics: Methods and Models.* The Netherlands: Kluwer Academic Publishers.

Anselin L. (2001). Rao's score test in spatial econometrics. *Journal of Statistical Planning and Inferences* **97**, 113-139.

[3] Anselin L. and 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.

[4] Bao, Y. (2010). Finite sample bias of QMLE in spatial autoregressive models. Paper presented in the *IVth World Conference of the Spatial Econometrics Association*, June 9-12, 2010, Chicago (`http://www.agecon.purdue.edu/sea_2010/`).

[5] Bao, Y. and Ullah, A. (2007a) Finite sample properties of maximum likelihood estimator in spatial models. *Journal of Econometrics* **137**, 396-413.

[6] Bao, Y. and Ullah, A. (2007b) The second-order bias and mean squared error of estimators in time-series models. *Journal of Econometrics* **140**, 650-669.

[7] Besley, T. and Case, A. C. (1995). Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition. *The American Economic Review* **85**, 25-45.

[8] Bell, K. P. and Bockstael, N. E. (2000). Applying the generalized-moment estimation approach to spatial problems involving microlevel data. *Review of Economics and Statistics* **82**, 72-82.

[9] Bertrand, M., Luttmer, E. F. P. and Mullaninathan, S. (2000). Network effects and welfare cultures. *Quarterly Journal of Economics*, **115**, 1019-1055.

[10] Brueckner, J. K. (1998). Testing for Strategic Interaction Among Local Governments: The Case of Growth Controls. *Journal of Urban Economics* **44**, 438-467.

[11] Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society* **B 26**, 211-246.

[12] Case, A. C.(1991). Spatial Patterns in Household Demand. *Econometrica* **59**, 953-965.

[13] Case, A. C., Rosen, H. S. and Hines, J. R. (1993). Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States. *Journal of Public Economics* **52**, 285-307.

[14] Cliff, A. D. and Ord, J. K. (1973). *Spatial Autocorrelation.* London: Pion.

[15] Cliff, A. D. and Ord, J. K. (1981). *Spatial Process, Models and Applications.* London: Pion.

[16] Cressie, N. (1993). *Statistics for Spatial Data.* New York: John Wiley & Sons.

[17] Efron, B. (1979). Bootstrap method: another look at the jackknife. *The Annals of Statistics* **7**, 1-26.

[18] Hahn J. and Kuersteiner (2002). Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large. *Econometrica* **70**, 1639-1657.

[19] Hahn J. and Newey, W. (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica* **72**, 1295-1319.

[20] Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics* **35**, 73-101.

[21] Kelejian H. H. and 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.

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

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

[24] Laskar, M. R. and King, M. L. (1998). Modified likelihood and related methods for handling nuisance parameters in the linear regression model. *Monash Econometrics and Business Statistics Working Papers, no. 5/98.*

[25] Lee, L. F. (2003). Consistency and efficiency of least squares estimation for mixed regressive spatial autoregressive models. *Econometric Theory* **18**, 252-277.

[26] Lee, L. F. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive models with autoregressive disturbances. *Econometric Reviews* **22**, 307-335.

[27] Lee, L. F. (2004a). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **72**, 1899-1925.

[28] Lee, L. F. (2004b). A supplement to 'Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models'. Working paper, Department of Economics, Ohio State University.

[29] Lee, L. F. (2007a). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics* **137**, 489-514.

[30] Lee, L. F. (2007b). Identification and estimation of econometric models with group interaction, contextual factors and fixed effects. *Journal of Econometrics* **140**, 333-374.

[31] Lahiri, S. N. (2003). *Resampling Methods for Dependent Data.* New York: Springer.

[32] Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **37**, 17-33.

[33] Mynbaev, K. T. and Ullah, A. (2008). Asymptotic distribution of the OLS estimator for a purely autoregressive spatial model. *Journal of Multivariate Analysis* **99**, 245-277.

[34] Pace, R. K. ad Barry, R. P. (1997). Quick computation of spatial autoregressive estimators. *Geographical Analysis* **29**, 232-246.

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

[36] Rilstone, P., Srivastava, V. K. and Ullah, A. (1996). The second-order bias and mean squared error of nonlinear estimators. *Journal of Econometrics* **75**, 369-395.

[37] Robinson, P. M. (2010). Efficient estimation of the semiparametric spatial autoregressive model. *Journal of Econometrics* **157**, 6-17.

[38] Smirnov, O. and Anselin, L. (2001). Fast maximum likelihood estimation of very large spatial autoregressive models: a characteristic polynomial approach. *Computational Statistics and Data Analysis* **35**, 301-319.

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

[40] Ullah, A. (2004). *Finite Sample Econometrics.* New York: Oxford University Press.

[41] Ullah, A. and Srivastava, V. K. (1994). Moments of the ratio of quadratic forms in non-normal variables with econometric examples. *Journal of Econometrics* **62**, 129-141.

[42] 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) of the (Q)MLE of λ: Queen Contiguity
`MRSAR-A:` $\sigma = 3$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{ba}}$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\lambda}_n^{\text{bc3}}$
0.5	50	1	0.418 [.187](.168)	0.396 [.198](.168)	0.488 [.168](.168)	0.488 [.167](.167)
		2	0.423 [.178](.161)	0.402 [.189](.161)	0.487 [.160](.160)	0.487 [.160](.159)
		3	0.431 [.166](.150)	0.411 [.176](.151)	0.486 [.150](.149)	0.486 [.150](.149)
	100	1	0.465 [.107](.101)	0.457 [.110](.102)	0.495 [.101](.101)	0.495 [.101](.101)
		2	0.466 [.104](.099)	0.459 [.108](.099)	0.494 [.099](.099)	0.494 [.099](.099)
		3	0.470 [.096](.091)	0.463 [.099](.092)	0.494 [.091](.090)	0.493 [.091](.090)
	200	1	0.484 [.073](.071)	0.481 [.074](.071)	0.499 [.071](.071)	0.499 [.071](.071)
		2	0.483 [.072](.070)	0.481 [.073](.070)	0.498 [.070](.070)	0.498 [.070](.070)
		3	0.485 [.067](.065)	0.483 [.068](.065)	0.498 [.065](.065)	0.498 [.065](.065)
	500	1	0.494 [.046](.046)	0.493 [.046](.046)	0.501 [.046](.046)	0.501 [.046](.046)
		2	0.493 [.046](.046)	0.492 [.046](.046)	0.500 [.046](.046)	0.500 [.046](.046)
		3	0.493 [.044](.044)	0.492 [.044](.044)	0.499 [.044](.044)	0.499 [.044](.044)
0.2	50	1	0.144 [.172](.162)	0.130 [.177](.162)	0.191 [.163](.163)	0.192 [.163](.162)
		2	0.150 [.163](.155)	0.137 [.168](.156)	0.193 [.156](.156)	0.194 [.155](.155)
		3	0.156 [.152](.145)	0.143 [.157](.146)	0.190 [.145](.144)	0.191 [.145](.144)
	100	1	0.170 [.126](.123)	0.166 [.128](.123)	0.198 [.123](.123)	0.198 [.123](.123)
		2	0.171 [.121](.118)	0.167 [.123](.118)	0.197 [.118](.118)	0.196 [.118](.118)
		3	0.173 [.112](.109)	0.170 [.113](.109)	0.196 [.109](.109)	0.195 [.109](.109)
	200	1	0.183 [.093](.091)	0.179 [.094](.092)	0.200 [.092](.092)	0.200 [.092](.092)
		2	0.183 [.090](.088)	0.179 [.091](.089)	0.200 [.089](.089)	0.200 [.088](.088)
		3	0.183 [.086](.085)	0.180 [.087](.085)	0.199 [.085](.085)	0.198 [.085](.085)
	500	1	0.194 [.055](.055)	0.193 [.055](.055)	0.200 [.055](.055)	0.200 [.055](.055)
		2	0.194 [.054](.054)	0.193 [.055](.054)	0.200 [.054](.054)	0.200 [.054](.054)
		3	0.195 [.053](.052)	0.194 [.053](.053)	0.200 [.053](.053)	0.200 [.053](.053)
-0.2	50	1	-0.248 [.199](.193)	-0.255 [.201](.193)	-0.199 [.197](.197)	-0.197 [.196](.196)
		2	-0.244 [.194](.189)	-0.251 [.196](.189)	-0.201 [.192](.192)	-0.199 [.191](.191)
		3	-0.241 [.167](.162)	-0.247 [.169](.162)	-0.205 [.164](.164)	-0.206 [.163](.163)
	100	1	-0.221 [.133](.132)	-0.226 [.134](.132)	-0.198 [.133](.133)	-0.198 [.133](.133)
		2	-0.222 [.129](.127)	-0.228 [.131](.128)	-0.202 [.129](.129)	-0.201 [.129](.129)
			-0.217 [.121](.120)	-0.222 [.122](.120)	-0.199 [.120](.120)	-0.199 [.120](.120)
	200	1	-0.214 [.105](.104)	-0.219 [.106](.104)	-0.201 [.105](.105)	-0.200 [.105](.105)
		2	-0.214 [.103](.102)	-0.218 [.104](.102)	-0.201 [.103](.103)	-0.201 [.103](.103)
		3	-0.214 [.097](.096)	-0.218 [.098](.096)	-0.202 [.097](.097)	-0.202 [.097](.097)
	500	1	-0.206 [.065](.065)	-0.207 [.065](.065)	-0.200 [.065](.065)	-0.200 [.065](.065)
		2	-0.206 [.065](.064)	-0.207 [.065](.065)	-0.200 [.065](.065)	-0.200 [.065](.065)
		3	-0.204 [.061](.061)	-0.205 [.061](.061)	-0.199 [.061](.061)	-0.199 [.061](.061)
-0.5	50	1	-0.531 [.217](.215)	-0.543 [.221](.216)	-0.489 [.223](.223)	-0.483 [.222](.221)
		2	-0.531 [.206](.204)	-0.543 [.210](.205)	-0.493 [.210](.210)	-0.488 [.209](.209)
		3	-0.527 [.187](.185)	-0.537 [.190](.186)	-0.495 [.191](.190)	-0.490 [.189](.189)
	100	1	-0.523 [.148](.146)	-0.532 [.150](.147)	-0.495 [.149](.148)	-0.494 [.148](.148)
		2	-0.523 [.145](.143)	-0.532 [.147](.143)	-0.497 [.145](.145)	-0.496 [.145](.145)
		3	-0.520 [.133](.131)	-0.528 [.135](.132)	-0.498 [.133](.133)	-0.497 [.133](.133)
	200	1	-0.511 [.108](.108)	-0.515 [.109](.108)	-0.498 [.109](.109)	-0.498 [.109](.109)
		2	-0.509 [.107](.106)	-0.513 [.107](.107)	-0.497 [.107](.107)	-0.497 [.107](.107)
		3	-0.510 [.102](.101)	-0.513 [.103](.102)	-0.499 [.102](.102)	-0.499 [.102](.102)
	500	1	-0.503 [.065](.064)	-0.504 [.065](.065)	-0.498 [.065](.065)	-0.498 [.065](.065)
		2	-0.504 [.064](.064)	-0.505 [.065](.064)	-0.499 [.064](.064)	-0.499 [.064](.064)
		3	-0.504 [.063](.063)	-0.505 [.063](.063)	-0.499 [.063](.063)	-0.499 [.063](.063)

Table 2. Empirical Mean[rmse](sd) of the (Q)MLE of λ: Group Interaction with $k = n^{0.5}$
MRSAR-B: $\sigma = 1$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{ba}}$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\lambda}_n^{\text{bc3}}$
0.5	50	1	0.482 [.062](.060)	0.471 [.066](.060)	0.498 [.060](.060)	0.499 [.060](.060)
		2	0.483 [.062](.060)	0.473 [.066](.060)	0.496 [.060](.060)	0.498 [.060](.060)
		3	0.485 [.057](.055)	0.476 [.061](.056)	0.495 [.055](.055)	0.498 [.055](.055)
	100	1	0.492 [.042](.042)	0.484 [.044](.041)	0.500 [.042](.042)	0.500 [.042](.042)
		2	0.491 [.042](.041)	0.485 [.044](.041)	0.499 [.042](.041)	0.499 [.042](.042)
		3	0.492 [.041](.040)	0.486 [.043](.040)	0.498 [.040](.040)	0.500 [.040](.040)
	200	1	0.495 [.032](.032)	0.491 [.033](.032)	0.500 [.032](.032)	0.500 [.032](.032)
		2	0.495 [.032](.031)	0.491 [.033](.031)	0.499 [.032](.032)	0.499 [.032](.032)
		3	0.495 [.031](.031)	0.492 [.032](.031)	0.499 [.031](.031)	0.500 [.031](.031)
	500	1	0.498 [.021](.021)	0.496 [.021](.021)	0.500 [.021](.021)	0.500 [.021](.021)
		2	0.498 [.021](.021)	0.496 [.021](.021)	0.500 [.021](.021)	0.500 [.021](.021)
		3	0.498 [.021](.021)	0.496 [.021](.021)	0.500 [.021](.021)	0.500 [.021](.021)
0.2	50	1	0.173 [.103](.099)	0.148 [.111](.098)	0.201 [.102](.102)	0.204 [.102](.102)
		2	0.173 [.101](.097)	0.151 [.108](.096)	0.198 [.099](.099)	0.203 [.099](.099)
		3	0.174 [.098](.094)	0.156 [.105](.095)	0.193 [.095](.094)	0.200 [.095](.095)
	100	1	0.183 [.076](.074)	0.170 [.079](.073)	0.201 [.074](.074)	0.201 [.074](.074)
		2	0.183 [.075](.073)	0.170 [.079](.073)	0.199 [.074](.074)	0.201 [.074](.074)
		3	0.182 [.072](.070)	0.171 [.076](.071)	0.196 [.070](.070)	0.198 [.070](.070)
	200	1	0.193 [.049](.048)	0.187 [.050](.048)	0.200 [.049](.049)	0.201 [.049](.049)
		2	0.192 [.049](.048)	0.186 [.050](.048)	0.199 [.049](.049)	0.199 [.049](.049)
		3	0.193 [.048](.047)	0.188 [.049](.047)	0.199 [.048](.048)	0.200 [.048](.048)
	500	1	0.197 [.032](.032)	0.194 [.032](.032)	0.200 [.032](.032)	0.200 [.032](.032)
		2	0.197 [.032](.032)	0.194 [.032](.032)	0.200 [.032](.032)	0.200 [.032](.032)
		3	0.196 [.032](.031)	0.194 [.032](.031)	0.199 [.031](.031)	0.200 [.031](.031)
-0.2	50	1	-0.240 [.143](.137)	-0.272 [.153](.135)	-0.198 [.142](.142)	-0.192 [.142](.142)
		2	-0.238 [.139](.134)	-0.267 [.148](.133)	-0.201 [.137](.137)	-0.192 [.137](.137)
		3	-0.237 [.130](.125)	-0.259 [.139](.126)	-0.207 [.126](.126)	-0.195 [.127](.127)
	100	1	-0.226 [.109](.106)	-0.245 [.114](.105)	-0.199 [.108](.108)	-0.197 [.108](.108)
		2	-0.226 [.109](.105)	-0.245 [.114](.105)	-0.201 [.107](.107)	-0.199 [.107](.107)
		3	-0.226 [.111](.108)	-0.241 [.116](.109)	-0.204 [.108](.108)	-0.200 [.108](.108)
	200	1	-0.211 [.075](.074)	-0.220 [.076](.074)	-0.200 [.075](.075)	-0.200 [.075](.075)
		2	-0.210 [.075](.074)	-0.218 [.076](.074)	-0.199 [.074](.074)	-0.199 [.074](.074)
		3	-0.211 [.075](.074)	-0.218 [.076](.074)	-0.201 [.075](.075)	-0.200 [.075](.075)
	500	1	-0.206 [.050](.049)	-0.210 [.050](.049)	-0.200 [.050](.050)	-0.200 [.050](.050)
		2	-0.206 [.049](.049)	-0.210 [.050](.049)	-0.200 [.049](.049)	-0.200 [.049](.049)
		3	-0.207 [.050](.050)	-0.211 [.051](.050)	-0.202 [.050](.050)	-0.202 [.050](.050)
-0.5	50	1	-0.548 [.175](.168)	-0.581 [.183](.164)	-0.494 [.173](.173)	-0.485 [.174](.173)
		2	-0.547 [.171](.165)	-0.576 [.179](.162)	-0.499 [.169](.169)	-0.489 [.170](.169)
		3	-0.538 [.160](.155)	-0.563 [.166](.154)	-0.501 [.157](.157)	-0.490 [.159](.159)
	100	1	-0.522 [.120](.118)	-0.539 [.123](.117)	-0.496 [.120](.120)	-0.494 [.120](.120)
		2	-0.522 [.119](.117)	-0.538 [.123](.117)	-0.498 [.119](.119)	-0.496 [.119](.119)
		3	-0.521 [.113](.111)	-0.535 [.117](.111)	-0.501 [.112](.112)	-0.498 [.113](.112)
	200	1	-0.514 [.089](.088)	-0.526 [.091](.087)	-0.499 [.088](.088)	-0.499 [.088](.088)
		2	-0.517 [.090](.088)	-0.528 [.092](.088)	-0.502 [.089](.089)	-0.501 [.089](.089)
		3	-0.514 [.086](.084)	-0.524 [.088](.084)	-0.501 [.085](.085)	-0.500 [.085](.085)
	500	1	-0.506 [.059](.059)	-0.511 [.060](.059)	-0.500 [.059](.059)	-0.499 [.059](.059)
		2	-0.507 [.058](.058)	-0.512 [.059](.058)	-0.500 [.058](.058)	-0.500 [.058](.058)
		3	-0.506 [.058](.058)	-0.511 [.059](.058)	-0.500 [.058](.058)	-0.500 [.058](.058)

Table 3. Empirical Mean[rmse](sd) of the (Q)MLE of λ: Group Interaction with $k = n^{0.5}$

`MRSAR-B:` $\sigma = 3$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{ba}}$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\lambda}_n^{\text{bc3}}$
0.5	50	1	0.398 [.174](.141)	0.350 [.210](.146)	0.490 [.137](.137)	0.495 [.136](.136)
		2	0.401 [.171](.139)	0.354 [.205](.144)	0.484 [.135](.134)	0.490 [.133](.133)
		3	0.414 [.156](.131)	0.373 [.187](.137)	0.485 [.124](.123)	0.493 [.121](.121)
	100	1	0.445 [.116](.102)	0.420 [.131](.104)	0.494 [.102](.101)	0.495 [.101](.101)
		2	0.447 [.111](.098)	0.423 [.127](.100)	0.494 [.098](.097)	0.495 [.097](.097)
		3	0.453 [.103](.091)	0.429 [.117](.093)	0.492 [.091](.090)	0.494 [.090](.090)
	200	1	0.470 [.077](.071)	0.453 [.086](.072)	0.497 [.072](.072)	0.497 [.072](.072)
		2	0.472 [.075](.069)	0.456 [.083](.070)	0.498 [.070](.070)	0.499 [.070](.070)
		3	0.475 [.072](.067)	0.459 [.079](.068)	0.498 [.067](.067)	0.499 [.067](.067)
	500	1	0.485 [.053](.051)	0.477 [.056](.051)	0.499 [.051](.051)	0.499 [.051](.051)
		2	0.487 [.051](.050)	0.479 [.054](.050)	0.500 [.050](.050)	0.500 [.050](.050)
		3	0.486 [.051](.049)	0.479 [.054](.049)	0.499 [.049](.049)	0.499 [.049](.049)
0.2	50	1	0.068 [.239](.199)	0.007 [.280](.202)	0.191 [.201](.200)	0.199 [.198](.198)
		2	0.077 [.225](.188)	0.018 [.264](.192)	0.189 [.189](.188)	0.198 [.186](.186)
		3	0.092 [.207](.176)	0.038 [.242](.180)	0.185 [.175](.174)	0.196 [.172](.171)
	100	1	0.123 [.168](.149)	0.086 [.190](.152)	0.198 [.150](.150)	0.200 [.150](.150)
		2	0.122 [.167](.148)	0.085 [.189](.150)	0.192 [.148](.148)	0.195 [.148](.148)
		3	0.130 [.152](.134)	0.095 [.173](.137)	0.190 [.135](.135)	0.192 [.134](.134)
	200	1	0.152 [.121](.111)	0.126 [.134](.112)	0.197 [.112](.112)	0.198 [.112](.112)
		2	0.153 [.120](.110)	0.127 [.133](.111)	0.196 [.111](.111)	0.197 [.111](.111)
		3	0.157 [.111](.103)	0.133 [.124](.104)	0.196 [.103](.103)	0.197 [.103](.103)
	500	1	0.178 [.082](.079)	0.166 [.087](.079)	0.200 [.080](.080)	0.200 [.080](.080)
		2	0.177 [.082](.079)	0.165 [.087](.079)	0.198 [.080](.080)	0.198 [.080](.080)
		3	0.179 [.080](.077)	0.167 [.084](.077)	0.199 [.077](.077)	0.199 [.077](.077)
-0.2	50	1	-0.386 [.340](.285)	-0.447 [.377](.284)	-0.200 [.278](.278)	-0.182 [.274](.273)
		2	-0.372 [.323](.273)	-0.432 [.359](.274)	-0.203 [.266](.266)	-0.186 [.261](.261)
		3	-0.344 [.285](.246)	-0.400 [.318](.248)	-0.203 [.240](.240)	-0.187 [.236](.235)
	100	1	-0.300 [.232](.209)	-0.343 [.255](.211)	-0.197 [.210](.210)	-0.193 [.209](.209)
		2	-0.301 [.230](.207)	-0.344 [.254](.209)	-0.204 [.208](.208)	-0.200 [.207](.207)
		3	-0.288 [.204](.184)	-0.329 [.227](.187)	-0.206 [.185](.185)	-0.203 [.184](.184)
	200	1	-0.275 [.184](.168)	-0.312 [.203](.169)	-0.204 [.169](.169)	-0.202 [.169](.169)
		2	-0.273 [.182](.166)	-0.309 [.200](.168)	-0.204 [.168](.168)	-0.203 [.168](.168)
		3	-0.265 [.170](.157)	-0.300 [.187](.159)	-0.204 [.158](.158)	-0.202 [.157](.157)
	500	1	-0.232 [.118](.113)	-0.250 [.124](.114)	-0.201 [.115](.115)	-0.200 [.115](.115)
		2	-0.233 [.117](.112)	-0.251 [.124](.113)	-0.202 [.114](.114)	-0.202 [.114](.113)
		3	-0.230 [.113](.109)	-0.247 [.119](.110)	-0.201 [.110](.110)	-0.201 [.110](.110)
-0.5	50	1	-0.723 [.382](.310)	-0.822 [.443](.304)	-0.480 [.323](.322)	-0.446 [.321](.316)
		2	-0.712 [.367](.299)	-0.809 [.427](.295)	-0.492 [.309](.309)	-0.456 [.303](.300)
		3	-0.689 [.335](.276)	-0.779 [.392](.275)	-0.502 [.281](.281)	-0.470 [.274](.273)
	100	1	-0.618 [.264](.236)	-0.678 [.296](.236)	-0.496 [.242](.242)	-0.490 [.241](.241)
		2	-0.613 [.259](.233)	-0.671 [.288](.232)	-0.500 [.237](.237)	-0.493 [.236](.236)
		3	-0.602 [.243](.220)	-0.655 [.270](.221)	-0.505 [.223](.223)	-0.497 [.222](.222)
	200	1	-0.595 [.235](.215)	-0.638 [.257](.216)	-0.500 [.216](.216)	-0.498 [.216](.216)
		2	-0.594 [.230](.210)	-0.637 [.252](.212)	-0.503 [.212](.212)	-0.501 [.212](.212)
		3	-0.588 [.221](.203)	-0.629 [.242](.205)	-0.505 [.204](.204)	-0.503 [.204](.204)
	500	1	-0.543 [.148](.142)	-0.567 [.157](.142)	-0.500 [.143](.143)	-0.500 [.143](.143)
		2	-0.543 [.149](.142)	-0.567 [.158](.143)	-0.501 [.144](.144)	-0.501 [.144](.144)
		3	-0.540 [.145](.139)	-0.564 [.154](.140)	-0.502 [.141](.141)	-0.501 [.141](.141)

Table 4. Empirical Mean[rmse](sd) of the (Q)MLE of λ: Group Interaction with $k = n^{0.35}$
MRSAR-B: $\sigma = 1$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{ba}}$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\lambda}_n^{\text{bc3}}$
0.5	50	1	0.477 [.071](.067)	0.460 [.077](.066)	0.498 [.069](.069)	0.500 [.069](.069)
		2	0.477 [.069](.065)	0.463 [.075](.065)	0.496 [.067](.066)	0.500 [.067](.067)
		3	0.479 [.070](.066)	0.468 [.075](.068)	0.494 [.066](.066)	0.500 [.067](.067)
	100	1	0.488 [.054](.053)	0.479 [.057](.053)	0.500 [.054](.054)	0.501 [.054](.054)
		2	0.487 [.054](.053)	0.478 [.057](.052)	0.498 [.053](.053)	0.499 [.053](.053)
		3	0.488 [.054](.052)	0.480 [.056](.053)	0.497 [.053](.053)	0.499 [.053](.053)
	200	1	0.495 [.036](.036)	0.490 [.037](.036)	0.500 [.036](.036)	0.500 [.036](.036)
		2	0.494 [.036](.036)	0.489 [.037](.036)	0.499 [.036](.036)	0.500 [.036](.036)
		3	0.494 [.036](.035)	0.490 [.036](.035)	0.499 [.035](.035)	0.500 [.036](.036)
	500	1	0.498 [.025](.025)	0.495 [.025](.025)	0.500 [.025](.025)	0.500 [.025](.025)
		2	0.497 [.025](.025)	0.495 [.025](.025)	0.500 [.025](.025)	0.500 [.025](.025)
		3	0.497 [.025](.025)	0.495 [.025](.025)	0.500 [.025](.025)	0.500 [.025](.025)
0.2	50	1	0.173 [.097](.093)	0.152 [.103](.092)	0.199 [.095](.095)	0.202 [.096](.096)
		2	0.175 [.096](.093)	0.157 [.102](.092)	0.198 [.095](.095)	0.203 [.096](.095)
		3	0.176 [.091](.088)	0.162 [.097](.089)	0.194 [.088](.088)	0.202 [.090](.090)
	100	1	0.176 [.091](.088)	0.166 [.094](.088)	0.202 [.088](.088)	0.203 [.088](.088)
		2	0.173 [.092](.088)	0.163 [.096](.088)	0.196 [.088](.088)	0.197 [.088](.088)
		3	0.174 [.102](.099)	0.166 [.106](.100)	0.194 [.097](.097)	0.195 [.097](.097)
	200	1	0.191 [.053](.052)	0.185 [.054](.052)	0.200 [.053](.053)	0.201 [.053](.053)
		2	0.192 [.053](.053)	0.185 [.055](.053)	0.200 [.053](.053)	0.201 [.053](.053)
		3	0.190 [.053](.052)	0.185 [.054](.052)	0.198 [.052](.052)	0.199 [.052](.052)
	500	1	0.196 [.037](.037)	0.193 [.038](.037)	0.200 [.037](.037)	0.200 [.037](.037)
		2	0.196 [.037](.037)	0.193 [.037](.037)	0.200 [.037](.037)	0.200 [.037](.037)
		3	0.196 [.037](.037)	0.193 [.037](.037)	0.200 [.037](.037)	0.200 [.037](.037)
-0.2	50	1	-0.245 [.162](.156)	-0.276 [.171](.153)	-0.199 [.161](.161)	-0.193 [.161](.161)
		2	-0.241 [.161](.155)	-0.270 [.169](.154)	-0.201 [.159](.159)	-0.193 [.160](.160)
		3	-0.238 [.156](.151)	-0.260 [.164](.152)	-0.206 [.152](.152)	-0.194 [.155](.155)
	100	1	-0.228 [.110](.106)	-0.249 [.116](.105)	-0.201 [.109](.109)	-0.199 [.109](.109)
		2	-0.226 [.109](.106)	-0.246 [.115](.105)	-0.201 [.107](.107)	-0.199 [.108](.108)
			-0.227 [.108](.104)	-0.243 [.113](.104)	-0.206 [.105](.105)	-0.201 [.106](.106)
	200	1	-0.212 [.079](.078)	-0.221 [.080](.078)	-0.201 [.079](.079)	-0.200 [.079](.079)
		2	-0.211 [.079](.078)	-0.219 [.080](.078)	-0.200 [.078](.078)	-0.200 [.078](.078)
		3	-0.211 [.076](.075)	-0.218 [.077](.075)	-0.202 [.076](.076)	-0.201 [.076](.076)
	500	1	-0.206 [.051](.051)	-0.211 [.052](.051)	-0.201 [.051](.051)	-0.201 [.051](.051)
		2	-0.205 [.051](.051)	-0.210 [.052](.051)	-0.199 [.051](.051)	-0.199 [.051](.051)
		3	-0.206 [.052](.052)	-0.210 [.053](.052)	-0.201 [.052](.052)	-0.201 [.052](.052)
-0.5	50	1	-0.571 [.208](.195)	-0.617 [.224](.191)	-0.495 [.202](.202)	-0.485 [.203](.203)
		2	-0.569 [.202](.189)	-0.609 [.217](.188)	-0.502 [.194](.194)	-0.490 [.196](.195)
		3	-0.569 [.212](.200)	-0.601 [.226](.202)	-0.515 [.199](.198)	-0.499 [.200](.200)
	100	1	-0.532 [.141](.138)	-0.555 [.147](.136)	-0.500 [.140](.140)	-0.498 [.140](.140)
		2	-0.530 [.139](.135)	-0.551 [.144](.134)	-0.500 [.138](.138)	-0.497 [.138](.138)
		3	-0.531 [.139](.135)	-0.549 [.144](.135)	-0.505 [.137](.137)	-0.500 [.137](.137)
	200	1	-0.516 [.092](.091)	-0.528 [.095](.091)	-0.501 [.092](.092)	-0.500 [.092](.092)
		2	-0.515 [.093](.091)	-0.527 [.095](.091)	-0.501 [.092](.092)	-0.500 [.092](.092)
		3	-0.515 [.092](.091)	-0.525 [.094](.091)	-0.502 [.091](.091)	-0.500 [.091](.091)
	500	1	-0.507 [.063](.063)	-0.512 [.064](.063)	-0.500 [.063](.063)	-0.500 [.063](.063)
		2	-0.507 [.064](.063)	-0.513 [.064](.063)	-0.501 [.064](.064)	-0.500 [.064](.064)
		3	-0.507 [.064](.064)	-0.512 [.065](.064)	-0.501 [.064](.064)	-0.500 [.064](.064)

Table 5. Empirical Mean[rmse](sd) of the (Q)MLE of λ: Group Interaction with $k = n^{0.35}$
MRSAR-B: $\sigma = 3$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{ba}}$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\lambda}_n^{\text{bc3}}$
0.5	50	1	0.331 [.254](.189)	0.258 [.313](.198)	0.478 [.196](.194)	0.485 [.193](.193)
		2	0.335 [.250](.188)	0.266 [.306](.197)	0.468 [.192](.190)	0.478 [.188](.187)
		3	0.360 [.223](.174)	0.299 [.272](.183)	0.468 [.175](.172)	0.480 [.170](.169)
	100	1	0.415 [.164](.140)	0.383 [.185](.143)	0.491 [.138](.137)	0.493 [.137](.137)
		2	0.417 [.162](.139)	0.387 [.182](.143)	0.488 [.136](.136)	0.490 [.135](.135)
		3	0.420 [.163](.141)	0.393 [.181](.147)	0.482 [.135](.134)	0.485 [.134](.134)
	200	1	0.448 [.112](.099)	0.422 [.127](.100)	0.497 [.102](.102)	0.497 [.102](.102)
		2	0.450 [.111](.099)	0.424 [.126](.101)	0.497 [.101](.101)	0.497 [.101](.101)
		3	0.451 [.107](.095)	0.426 [.122](.097)	0.491 [.097](.097)	0.492 [.096](.096)
	500	1	0.481 [.065](.062)	0.469 [.069](.062)	0.499 [.063](.063)	0.499 [.063](.063)
		2	0.481 [.065](.062)	0.469 [.069](.062)	0.499 [.063](.063)	0.499 [.063](.063)
		3	0.481 [.063](.060)	0.470 [.068](.061)	0.498 [.061](.061)	0.498 [.061](.061)
0.2	50	1	0.027 [.280](.221)	-0.064 [.343](.220)	0.186 [.233](.233)	0.201 [.232](.232)
		2	0.030 [.281](.223)	-0.054 [.340](.227)	0.172 [.230](.228)	0.191 [.226](.226)
		3	0.042 [.293](.247)	-0.027 [.342](.256)	0.158 [.242](.238)	0.181 [.235](.234)
	100	1	0.087 [.209](.176)	0.027 [.248](.177)	0.190 [.184](.184)	0.194 [.184](.184)
		2	0.093 [.205](.175)	0.035 [.242](.177)	0.189 [.181](.180)	0.194 [.180](.180)
		3	0.099 [.202](.174)	0.048 [.235](.179)	0.180 [.176](.175)	0.187 [.174](.174)
	200	1	0.134 [.157](.143)	0.101 [.175](.143)	0.198 [.146](.146)	0.199 [.146](.146)
		2	0.135 [.158](.143)	0.102 [.174](.145)	0.196 [.146](.146)	0.197 [.146](.146)
		3	0.134 [.158](.144)	0.105 [.174](.146)	0.189 [.145](.145)	0.191 [.144](.144)
	500	1	0.168 [.099](.093)	0.147 [.107](.093)	0.199 [.095](.095)	0.199 [.095](.095)
		2	0.167 [.099](.093)	0.147 [.107](.093)	0.197 [.095](.095)	0.198 [.095](.095)
		3	0.169 [.097](.092)	0.150 [.105](.092)	0.198 [.093](.093)	0.199 [.093](.093)
-0.2	50	1	-0.455 [.443](.362)	-0.550 [.505](.364)	-0.201 [.375](.375)	-0.179 [.370](.369)
		2	-0.453 [.457](.380)	-0.542 [.512](.381)	-0.226 [.385](.384)	-0.200 [.376](.376)
		3	-0.442 [.523](.463)	-0.512 [.547](.450)	-0.244 [.437](.435)	-0.205 [.419](.419)
	100	1	-0.362 [.310](.264)	-0.445 [.361](.265)	-0.208 [.276](.276)	-0.202 [.275](.275)
		2	-0.357 [.308](.265)	-0.436 [.356](.267)	-0.214 [.274](.273)	-0.206 [.272](.272)
		3	-0.342 [.290](.253)	-0.413 [.334](.258)	-0.220 [.258](.257)	-0.210 [.255](.255)
	200	1	-0.294 [.225](.205)	-0.351 [.253](.204)	-0.205 [.213](.213)	-0.204 [.213](.213)
		2	-0.291 [.223](.204)	-0.345 [.251](.204)	-0.205 [.211](.211)	-0.203 [.211](.211)
		3	-0.283 [.215](.198)	-0.334 [.240](.199)	-0.208 [.204](.204)	-0.204 [.203](.203)
	500	1	-0.237 [.133](.128)	-0.266 [.143](.127)	-0.198 [.131](.131)	-0.197 [.131](.131)
		2	-0.241 [.134](.128)	-0.269 [.144](.127)	-0.202 [.131](.131)	-0.202 [.131](.131)
		3	-0.238 [.133](.127)	-0.265 [.142](.126)	-0.203 [.129](.129)	-0.201 [.129](.129)
-0.5	50	1	-0.915 [.700](.564)	-0.999 [.747](.556)	-0.487 [.533](.533)	-0.448 [.516](.513)
		2	-0.912 [.711](.579)	-0.994 [.748](.561)	-0.517 [.533](.532)	-0.466 [.506](.505)
		3	-0.845 [.643](.543)	-0.918 [.664](.516)	-0.508 [.485](.485)	-0.453 [.472](.470)
	100	1	-0.699 [.375](.318)	-0.798 [.435](.317)	-0.506 [.332](.332)	-0.498 [.331](.331)
		2	-0.690 [.368](.315)	-0.785 [.426](.317)	-0.511 [.324](.324)	-0.501 [.323](.323)
		3	-0.671 [.343](.298)	-0.755 [.395](.302)	-0.520 [.303](.302)	-0.507 [.300](.300)
	200	1	-0.629 [.291](.260)	-0.706 [.331](.259)	-0.503 [.271](.271)	-0.501 [.271](.271)
		2	-0.631 [.289](.257)	-0.706 [.329](.256)	-0.511 [.267](.266)	-0.508 [.266](.266)
		3	-0.616 [.281](.256)	-0.684 [.316](.257)	-0.509 [.262](.262)	-0.503 [.261](.261)
	500	1	-0.549 [.178](.171)	-0.584 [.190](.171)	-0.495 [.175](.175)	-0.495 [.175](.175)
		2	-0.555 [.183](.174)	-0.589 [.195](.174)	-0.503 [.178](.178)	-0.502 [.178](.178)
		3	-0.551 [.175](.168)	-0.584 [.188](.168)	-0.503 [.171](.171)	-0.501 [.171](.171)

Table 6. Empirical Means for All Parameter Estimators: Group Interaction with $k = n^{0.35}$

MRSAR-A: $\sigma = 1$, $\beta = (5, 1, .5)'$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{bc2}$	$\hat{\beta}_{0n}$	$\hat{\beta}_{1n}$	$\hat{\beta}_{2n}$	$\hat{\beta}_{0n}^{bc}$	$\hat{\beta}_{1n}^{bc}$	$\hat{\beta}_{2n}^{bc}$	$\tilde{\sigma}_n^2$	$\tilde{\sigma}_n^{2bc}$
0.5	50	1	0.457	0.491	6.000	0.997	0.500	5.217	0.998	0.499	0.980	0.978
		2	0.460	0.490	5.939	0.997	0.500	5.242	0.998	0.499	0.978	0.976
		3	0.457	0.483	6.022	0.997	0.500	5.417	0.998	0.499	1.000	0.997
	100	1	0.465	0.497	5.921	0.999	0.500	5.097	0.999	0.500	0.991	0.989
		2	0.465	0.495	5.899	1.000	0.499	5.133	0.999	0.499	0.988	0.986
		2	0.471	0.496	5.751	1.000	0.499	5.106	0.999	0.500	1.004	1.002
	200	1	0.489	0.499	5.303	1.000	0.500	5.031	1.000	0.500	0.994	0.994
		2	0.489	0.499	5.288	1.000	0.500	5.027	1.000	0.500	0.995	0.995
		2	0.489	0.498	5.277	1.000	0.500	5.046	1.000	0.500	0.988	0.988
	500	1	0.479	0.499	5.524	1.000	0.500	5.028	1.000	0.500	0.999	0.998
		2	0.480	0.500	5.500	1.000	0.500	5.014	1.000	0.500	0.997	0.996
		2	0.480	0.498	5.487	1.000	0.500	5.041	1.000	0.500	0.991	0.991
0.2	50	1	0.158	0.192	5.671	0.998	0.499	5.128	0.999	0.499	0.979	0.980
		2	0.159	0.190	5.655	0.998	0.499	5.170	0.998	0.499	0.975	0.976
		2	0.164	0.188	5.585	0.998	0.499	5.196	0.998	0.499	0.946	0.947
	100	1	0.169	0.198	5.476	1.000	0.500	5.044	0.999	0.500	0.989	0.989
		2	0.170	0.196	5.473	1.000	0.500	5.073	1.000	0.500	0.987	0.988
		2	0.174	0.196	5.402	1.000	0.500	5.069	0.999	0.500	0.982	0.982
	200	1	0.144	0.197	5.900	0.999	0.499	5.058	0.999	0.500	0.993	0.993
		2	0.146	0.197	5.865	0.999	0.499	5.057	1.000	0.500	0.997	0.997
		2	0.144	0.188	5.904	0.999	0.499	5.189	0.999	0.500	1.000	1.000
	500	1	0.189	0.200	5.167	1.000	0.500	5.001	1.000	0.500	0.998	0.998
		2	0.191	0.201	5.142	1.000	0.500	4.981	1.000	0.500	0.998	0.998
		2	0.189	0.199	5.172	1.000	0.500	5.021	1.000	0.500	0.994	0.994
-0.2	50	1	-0.347	-0.212	6.654	0.988	0.495	5.141	0.998	0.499	0.964	0.975
		2	-0.346	-0.225	6.644	0.988	0.494	5.292	0.997	0.499	0.962	0.972
		2	-0.338	-0.239	6.558	0.989	0.495	5.448	0.996	0.498	0.964	0.975
	100	1	-0.235	-0.200	5.362	0.998	0.500	5.006	0.999	0.500	0.987	0.989
		2	-0.237	-0.204	5.378	0.999	0.500	5.047	1.000	0.500	0.988	0.990
		2	-0.236	-0.208	5.371	0.999	0.500	5.090	1.000	0.499	1.006	1.008
	200	1	-0.248	-0.203	5.519	0.999	0.499	5.036	0.999	0.500	0.993	0.994
		2	-0.244	-0.201	5.478	0.999	0.500	5.016	0.999	0.500	0.992	0.993
		2	-0.244	-0.206	5.478	0.999	0.500	5.067	0.999	0.500	0.996	0.997
	500	1	-0.225	-0.200	5.265	1.000	0.500	5.000	1.000	0.500	0.997	0.998
		2	-0.226	-0.201	5.269	1.000	0.500	5.009	1.000	0.500	0.999	0.999
		2	-0.224	-0.201	5.251	1.000	0.500	5.009	1.000	0.500	0.996	0.996
-0.5	50	1	-0.538	-0.500	5.338	1.000	0.498	5.004	0.998	0.500	0.976	0.980
		2	-0.541	-0.506	5.358	0.999	0.498	5.062	0.998	0.500	0.971	0.975
		2	-0.535	-0.507	5.307	1.001	0.499	5.068	1.000	0.500	0.964	0.968
	100	1	-0.546	-0.502	5.397	1.000	0.499	5.025	1.000	0.500	0.988	0.990
		2	-0.544	-0.503	5.376	0.999	0.500	5.031	0.999	0.500	0.988	0.990
		2	-0.542	-0.508	5.352	1.000	0.500	5.066	1.000	0.500	0.979	0.981
	200	1	-0.529	-0.500	5.238	1.000	0.500	5.006	0.999	0.500	0.994	0.995
		2	-0.527	-0.500	5.225	1.000	0.500	5.004	0.999	0.500	0.992	0.993
		2	-0.530	-0.506	5.250	1.000	0.500	5.053	0.999	0.500	1.009	1.010
	500	1	-0.519	-0.500	5.165	1.000	0.500	5.002	1.000	0.500	0.998	0.998
		2	-0.518	-0.499	5.152	1.000	0.500	4.992	1.000	0.500	0.997	0.998
		2	-0.518	-0.501	5.155	1.000	0.500	5.007	1.000	0.500	0.994	0.995

Table 7. Empirical Means for All Parameter Estimators: Group Interaction with $k = n^{0.5}$

`MRSAR-A`: $\sigma = 3$, $\beta = (5, 1, .5)'$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\beta}_{0n}$	$\hat{\beta}_{1n}$	$\hat{\beta}_{2n}$	$\hat{\beta}_{0n}^{\text{bc}}$	$\hat{\beta}_{1n}^{\text{bc}}$	$\hat{\beta}_{2n}^{\text{bc}}$	$\tilde{\sigma}_n^2$	$\tilde{\sigma}_n^{2\text{bc}}$
0.5	50	1	0.440	0.488	6.462	1.002	0.500	5.325	0.993	0.499	8.874	8.795
		2	0.445	0.488	6.374	0.999	0.500	5.345	0.991	0.498	8.858	8.780
		3	0.451	0.487	6.203	1.004	0.498	5.353	0.997	0.497	8.749	8.675
	100	1	0.458	0.497	6.042	1.002	0.498	5.112	0.997	0.498	8.924	8.873
		2	0.458	0.494	6.042	1.001	0.500	5.161	0.997	0.500	8.984	8.933
		3	0.459	0.492	6.021	1.002	0.499	5.205	0.998	0.499	8.943	8.893
	200	1	0.468	0.498	5.853	0.997	0.500	5.072	0.998	0.499	8.977	8.947
		2	0.468	0.498	5.832	0.997	0.500	5.075	0.998	0.499	8.968	8.938
		3	0.470	0.497	5.783	0.999	0.500	5.088	1.000	0.499	9.044	9.015
	500	1	0.482	0.498	5.441	1.001	0.500	5.048	1.000	0.500	8.982	8.970
		2	0.484	0.500	5.383	1.000	0.500	4.998	0.999	0.500	9.014	9.002
		3	0.485	0.499	5.378	1.000	0.500	5.019	1.000	0.500	9.003	8.992
0.2	50	1	0.119	0.191	6.246	0.991	0.502	5.182	0.992	0.495	8.755	8.754
		2	0.125	0.190	6.136	0.996	0.502	5.180	0.997	0.496	8.871	8.868
		3	0.134	0.188	6.015	0.992	0.503	5.223	0.992	0.498	8.568	8.562
	100	1	0.130	0.197	6.173	0.993	0.498	5.084	0.997	0.498	8.891	8.883
		2	0.129	0.192	6.189	0.995	0.498	5.157	0.998	0.498	8.914	8.906
		3	0.144	0.197	5.955	0.994	0.499	5.073	0.996	0.499	8.865	8.855
	200	1	0.161	0.199	5.587	1.000	0.500	5.025	0.998	0.500	8.954	8.948
		2	0.161	0.197	5.602	0.998	0.500	5.059	0.997	0.500	8.996	8.990
		3	0.164	0.197	5.551	1.000	0.499	5.056	0.998	0.499	8.991	8.985
	500	1	0.172	0.200	5.441	0.999	0.500	5.012	0.999	0.500	8.984	8.980
		2	0.172	0.200	5.432	1.000	0.500	5.008	1.000	0.499	8.995	8.991
		2	0.174	0.199	5.409	1.000	0.500	5.012	1.000	0.500	8.963	8.960
-0.2	50	1	-0.346	-0.207	6.680	0.981	0.492	5.154	0.990	0.494	8.620	8.761
		2	-0.334	-0.207	6.537	0.984	0.493	5.145	0.992	0.495	8.681	8.812
		3	-0.312	-0.202	6.289	0.986	0.495	5.085	0.992	0.497	8.592	8.702
	100	1	-0.324	-0.203	6.373	0.988	0.493	5.055	0.997	0.498	8.804	8.879
		2	-0.317	-0.202	6.263	0.992	0.495	5.012	1.001	0.500	8.793	8.863
		3	-0.311	-0.205	6.217	0.991	0.494	5.065	0.999	0.499	8.854	8.918
	200	1	-0.281	-0.201	5.877	0.996	0.498	5.004	1.000	0.500	8.904	8.933
		2	-0.281	-0.203	5.880	0.995	0.498	5.032	0.999	0.500	8.908	8.936
		3	-0.270	-0.198	5.766	0.996	0.497	4.985	1.000	0.499	8.865	8.890
	500	1	-0.234	-0.199	5.358	1.001	0.500	4.988	1.000	0.500	8.983	8.989
		2	-0.233	-0.199	5.353	1.000	0.500	4.989	0.999	0.500	8.973	8.980
		3	-0.235	-0.202	5.370	1.000	0.499	5.027	0.999	0.499	8.981	8.988
-0.5	50	1	-0.648	-0.489	6.567	0.976	0.489	4.971	0.992	0.496	8.580	8.820
		2	-0.637	-0.493	6.418	0.981	0.491	4.969	0.996	0.497	8.610	8.835
		3	-0.618	-0.501	6.238	0.984	0.492	5.051	0.996	0.498	8.584	8.773
	100	1	-0.619	-0.498	6.093	0.988	0.493	5.000	0.997	0.499	8.778	8.890
		2	-0.609	-0.495	5.995	0.989	0.493	4.969	0.998	0.499	8.761	8.867
		3	-0.601	-0.498	5.941	0.988	0.494	5.011	0.996	0.499	8.818	8.915
	200	1	-0.577	-0.497	5.701	0.995	0.497	4.995	0.997	0.499	8.900	8.946
		2	-0.579	-0.502	5.713	0.995	0.497	5.033	0.997	0.499	8.900	8.945
		3	-0.566	-0.497	5.595	0.996	0.497	4.985	0.998	0.499	8.866	8.907
	500	1	-0.558	-0.502	5.501	0.998	0.499	5.025	0.999	0.500	8.965	8.983
		2	-0.556	-0.502	5.486	0.999	0.499	5.018	1.000	0.500	8.977	8.995
		3	-0.552	-0.501	5.454	0.998	0.499	5.017	0.999	0.500	8.932	8.949

Table 8. Empirical Means for All Parameter Estimators: Group Interaction with $k = n^{0.5}$

`MRSAR-B`: $\sigma = 1$, $\beta = (5, 1, .5)'$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\beta}_{0n}$	$\hat{\beta}_{1n}$	$\hat{\beta}_{2n}$	$\hat{\beta}_{0n}^{\text{bc}}$	$\hat{\beta}_{1n}^{\text{bc}}$	$\hat{\beta}_{2n}^{\text{bc}}$	$\tilde{\sigma}_n^2$	$\tilde{\sigma}_n^{2\text{bc}}$
0.5	50	1	0.482	0.499	5.194	1.032	0.507	5.014	1.001	0.499	0.979	0.978
		2	0.482	0.497	5.195	1.031	0.506	5.035	1.004	0.499	0.979	0.978
		3	0.483	0.495	5.192	1.029	0.508	5.062	1.008	0.501	1.006	1.004
	100	1	0.489	0.500	5.109	1.021	0.504	5.002	1.000	0.499	0.990	0.989
		2	0.489	0.499	5.112	1.021	0.503	5.013	1.002	0.499	0.987	0.986
		3	0.490	0.498	5.103	1.020	0.504	5.022	1.004	0.500	0.991	0.990
	200	1	0.495	0.500	5.053	1.010	0.503	5.002	1.000	0.500	0.995	0.995
		2	0.495	0.500	5.054	1.010	0.503	5.005	1.000	0.500	0.995	0.995
		3	0.496	0.500	5.046	1.008	0.503	5.002	1.000	0.500	0.994	0.993
	500	1	0.498	0.500	5.024	1.005	0.501	5.003	1.001	0.500	0.998	0.998
		2	0.498	0.500	5.021	1.004	0.501	5.000	1.000	0.500	0.999	0.999
		3	0.498	0.500	5.022	1.004	0.501	5.002	1.000	0.500	1.006	1.006
0.2	50	1	0.162	0.197	5.199	1.032	0.490	5.010	1.000	0.496	0.977	0.978
		2	0.163	0.194	5.195	1.030	0.492	5.028	1.002	0.497	0.979	0.979
		3	0.167	0.192	5.176	1.027	0.492	5.040	1.005	0.497	0.977	0.978
	100	1	0.183	0.200	5.111	1.021	0.502	5.005	1.001	0.500	0.990	0.990
		2	0.184	0.199	5.104	1.021	0.502	5.006	1.002	0.499	0.993	0.993
		3	0.183	0.196	5.111	1.021	0.503	5.028	1.005	0.500	1.012	1.012
	200	1	0.189	0.199	5.071	1.012	0.503	5.006	1.001	0.500	0.995	0.995
		2	0.191	0.200	5.064	1.011	0.501	5.002	1.000	0.499	0.994	0.994
		3	0.190	0.198	5.067	1.011	0.502	5.013	1.002	0.500	0.989	0.989
	500	1	0.196	0.200	5.026	1.005	0.501	5.001	1.000	0.500	0.999	0.999
		2	0.196	0.200	5.027	1.005	0.502	5.002	1.000	0.500	1.000	1.000
		3	0.196	0.200	5.024	1.004	0.501	5.001	1.000	0.500	0.997	0.997
-0.2	50	1	-0.259	-0.198	5.228	1.047	0.510	4.990	0.998	0.496	0.976	0.981
		2	-0.261	-0.209	5.237	1.049	0.511	5.030	1.007	0.500	0.967	0.971
		3	-0.253	-0.212	5.211	1.043	0.508	5.044	1.010	0.499	0.971	0.975
	100	1	-0.223	-0.197	5.104	1.017	0.500	4.990	0.997	0.500	0.989	0.990
		2	-0.225	-0.202	5.115	1.020	0.499	5.009	1.001	0.499	0.991	0.991
		3	-0.224	-0.205	5.110	1.019	0.501	5.023	1.003	0.500	0.985	0.985
	200	1	-0.212	-0.199	5.049	1.010	0.503	4.994	0.999	0.499	0.994	0.995
		2	-0.214	-0.201	5.055	1.011	0.503	5.003	1.000	0.500	0.994	0.994
		3	-0.213	-0.202	5.054	1.011	0.503	5.008	1.002	0.500	0.991	0.991
	500	1	-0.206	-0.200	5.025	1.004	0.501	5.002	1.000	0.500	0.998	0.998
		2	-0.205	-0.200	5.022	1.004	0.501	5.000	1.000	0.500	0.999	0.999
		3	-0.206	-0.201	5.024	1.004	0.501	5.003	1.001	0.500	0.998	0.998
-0.5	50	1	-0.543	-0.500	5.139	1.015	0.512	5.000	0.996	0.498	0.973	0.977
		2	-0.540	-0.503	5.134	1.015	0.511	5.011	0.998	0.499	0.972	0.976
		3	-0.535	-0.504	5.117	1.014	0.508	5.015	1.000	0.498	0.982	0.987
	100	1	-0.530	-0.499	5.099	1.019	0.506	4.995	0.999	0.499	0.987	0.989
		2	-0.529	-0.501	5.097	1.019	0.505	5.002	1.000	0.499	0.975	0.977
		3	-0.529	-0.505	5.098	1.018	0.506	5.018	1.003	0.501	0.986	0.988
	200	1	-0.517	-0.498	5.060	1.010	0.500	4.996	0.999	0.499	0.995	0.995
		2	-0.518	-0.501	5.065	1.011	0.499	5.003	1.000	0.499	0.991	0.991
		3	-0.517	-0.501	5.057	1.010	0.500	5.003	1.001	0.500	0.979	0.979
	500	1	-0.507	-0.501	5.024	1.005	0.500	5.003	1.001	0.500	0.997	0.997
		2	-0.507	-0.501	5.023	1.004	0.500	5.003	1.000	0.500	0.997	0.997
		3	-0.506	-0.500	5.021	1.004	0.500	5.002	1.000	0.500	0.994	0.994

Table 9. Empirical Means for All Parameter Estimators: Group Interaction with $k = n^{0.5}$

`MRSAR-B`: $\sigma = 3$, $\beta = (5, 1, .5)'$; dgp: 1=normal, 2=normal mixture($\tau = 4, p = .1$), 3=lognormal

λ	n	dgp	$\hat{\lambda}_n$	$\hat{\lambda}_n^{\text{bc2}}$	$\hat{\beta}_{0n}$	$\hat{\beta}_{1n}$	$\hat{\beta}_{2n}$	$\hat{\beta}_{0n}^{\text{bc}}$	$\hat{\beta}_{1n}^{\text{bc}}$	$\hat{\beta}_{2n}^{\text{bc}}$	$\tilde{\sigma}_n^2$	$\tilde{\sigma}_n^{2\text{bc}}$
0.5	50	1	0.409	0.491	5.896	1.172	0.566	5.085	1.015	0.503	8.894	8.778
		2	0.414	0.487	5.851	1.163	0.564	5.120	1.022	0.507	8.893	8.779
		3	0.421	0.482	5.819	1.150	0.560	5.190	1.032	0.512	9.037	8.926
	100	1	0.443	0.495	5.583	1.110	0.531	5.061	1.010	0.499	8.964	8.899
		2	0.445	0.493	5.550	1.107	0.535	5.063	1.013	0.504	8.925	8.861
		3	0.450	0.492	5.516	1.096	0.528	5.086	1.015	0.502	9.017	8.954
	200	1	0.473	0.499	5.271	1.048	0.507	5.004	0.999	0.496	8.951	8.923
		2	0.472	0.497	5.286	1.051	0.510	5.029	1.004	0.500	8.989	8.962
		3	0.473	0.496	5.277	1.048	0.510	5.043	1.006	0.501	8.974	8.947
	500	1	0.486	0.499	5.135	1.026	0.506	5.006	1.001	0.500	8.996	8.987
		2	0.487	0.499	5.132	1.026	0.503	5.005	1.001	0.497	9.001	8.992
		3	0.487	0.499	5.132	1.025	0.504	5.011	1.002	0.499	8.990	8.981
0.2	50	1	0.073	0.193	5.838	1.124	0.545	5.064	0.999	0.504	8.755	8.766
		2	0.082	0.190	5.760	1.118	0.532	5.066	1.005	0.496	8.736	8.745
		3	0.097	0.186	5.683	1.101	0.534	5.096	1.008	0.504	8.809	8.815
	100	1	0.122	0.198	5.472	1.091	0.508	5.008	1.001	0.490	8.906	8.898
		2	0.125	0.196	5.456	1.088	0.510	5.021	1.004	0.494	8.988	8.980
		3	0.131	0.193	5.422	1.079	0.510	5.041	1.007	0.496	8.894	8.886
	200	1	0.152	0.198	5.286	1.053	0.508	5.010	1.000	0.497	8.958	8.951
		2	0.153	0.197	5.279	1.052	0.508	5.014	1.001	0.498	8.957	8.950
		3	0.158	0.197	5.250	1.045	0.506	5.011	1.000	0.497	8.935	8.928
	500	1	0.175	0.199	5.153	1.030	0.506	5.006	1.001	0.500	8.969	8.966
		2	0.176	0.199	5.148	1.029	0.503	5.004	1.001	0.497	8.983	8.980
		3	0.176	0.198	5.149	1.028	0.504	5.015	1.002	0.499	8.976	8.973
-0.2	50	1	-0.357	-0.196	5.618	1.116	0.551	4.974	0.992	0.497	8.601	8.779
		2	-0.352	-0.207	5.585	1.113	0.547	5.009	1.002	0.498	8.659	8.830
		3	-0.331	-0.211	5.523	1.096	0.536	5.034	1.005	0.497	8.566	8.722
	100	1	-0.299	-0.199	5.414	1.072	0.505	4.991	0.997	0.492	8.823	8.883
		2	-0.297	-0.205	5.412	1.071	0.509	5.020	1.002	0.497	8.778	8.836
		3	-0.286	-0.206	5.366	1.063	0.511	5.024	1.003	0.500	8.769	8.823
	200	1	-0.274	-0.202	5.326	1.057	0.509	5.011	1.000	0.497	8.911	8.938
		2	-0.273	-0.203	5.320	1.054	0.507	5.018	1.000	0.495	8.874	8.900
		3	-0.268	-0.205	5.298	1.051	0.510	5.023	1.002	0.499	8.977	9.002
	500	1	-0.235	-0.201	5.148	1.028	0.509	5.003	1.000	0.500	8.968	8.974
		2	-0.233	-0.199	5.141	1.026	0.508	4.999	0.999	0.498	8.948	8.955
		3	-0.235	-0.203	5.149	1.027	0.509	5.016	1.002	0.500	8.943	8.949
-0.5	50	1	-0.719	-0.479	5.714	1.144	0.538	4.918	0.985	0.490	8.419	8.811
		2	-0.707	-0.490	5.682	1.137	0.532	4.964	0.994	0.489	8.522	8.900
		3	-0.681	-0.499	5.611	1.119	0.531	4.994	0.999	0.495	8.395	8.744
	100	1	-0.648	-0.492	5.447	1.086	0.512	4.962	0.992	0.490	8.750	8.902
		2	-0.643	-0.497	5.434	1.084	0.515	4.981	0.996	0.494	8.756	8.902
		3	-0.632	-0.502	5.410	1.075	0.514	5.002	0.997	0.496	8.854	8.998
	200	1	-0.592	-0.499	5.292	1.056	0.513	4.992	0.998	0.495	8.877	8.931
		2	-0.589	-0.500	5.285	1.055	0.514	4.997	0.999	0.498	8.886	8.940
		3	-0.580	-0.501	5.258	1.050	0.512	5.002	1.000	0.497	8.887	8.937
	500	1	-0.546	-0.500	5.153	1.029	0.507	5.001	1.000	0.498	8.969	8.984
		2	-0.546	-0.502	5.154	1.029	0.507	5.005	1.000	0.499	8.952	8.967
		3	-0.543	-0.502	5.147	1.027	0.508	5.009	1.001	0.500	9.048	9.062

Table 10. Performance of t-ratios: Group Interaction with $k = n^{0.35}$, `MRSAR-C`, $\sigma = 2$

Three rows under each n: `dgp1-dgp3`

		Each two columns for: $t(\lambda)$ and $t^{\mathrm{bc}}(\lambda)$						Each two columns for: $t(\beta_1)$ and $t^{\mathrm{bc}}(\beta_1)$					
λ	n	mean		sd		tail prob		mean		sd		tail prob	
0.5	50	-.752	.169	1.066	1.041	.130	.064	.386	-.089	1.131	1.105	.099	.079
		-.758	.091	1.014	1.030	.123	.057	.391	-.048	1.139	1.122	.106	.080
		-.700	.057	0.998	1.052	.105	.064	.328	-.071	1.105	1.091	.088	.073
	100	-.774	.109	1.016	1.026	.122	.057	.536	-.067	1.065	1.091	.098	.074
		-.730	.114	0.995	1.038	.111	.059	.473	-.097	1.057	1.081	.090	.072
		-.701	.058	0.966	1.049	.098	.061	.452	-.066	1.044	1.079	.084	.070
	200	-.659	.191	0.997	1.006	.100	.055	.156	-.085	1.071	1.024	.071	.058
		-.640	.189	0.991	1.019	.096	.060	.161	-.074	1.070	1.024	.071	.058
		-.623	.152	0.960	1.014	.087	.057	.158	-.065	1.063	1.025	.070	.057
	500	-.451	.131	1.013	1.032	.074	.058	.204	-.070	1.018	1.015	.057	.052
		-.456	.120	1.004	1.038	.075	.061	.207	-.062	1.034	1.030	.065	.059
		-.452	.099	0.975	1.048	.068	.063	.191	-.069	1.027	1.025	.061	.057
0.2	50	-.751	.218	1.099	1.043	.135	.069	.177	-.062	1.114	1.087	.080	.071
		-.717	.171	1.009	1.024	.105	.060	.197	-.034	1.095	1.077	.079	.070
		-.634	.159	0.948	0.973	.073	.050	.178	-.036	1.103	1.089	.080	.070
	100	-.702	.167	1.028	1.022	.113	.058	.332	-.118	1.099	1.078	.087	.072
		-.678	.145	1.005	1.042	.103	.063	.316	-.110	1.089	1.071	.084	.070
		-.652	.086	0.944	1.038	.082	.059	.332	-.063	1.059	1.057	.079	.065
	200	-.540	.086	1.003	1.028	.085	.057	.372	-.080	1.024	1.054	.069	.063
		-.549	.052	1.010	1.060	.091	.065	.391	-.044	1.028	1.060	.074	.064
		-.529	.017	0.991	1.073	.081	.072	.365	-.030	1.020	1.048	.071	.062
	500	-.436	.081	1.004	1.027	.069	.055	.325	-.063	1.009	1.035	.065	.059
		-.414	.096	1.003	1.036	.074	.058	.306	-.076	1.008	1.033	.061	.056
		-.410	.074	0.994	1.049	.069	.062	.289	-.071	1.007	1.030	.061	.056
-.2	50	-.975	.158	1.117	0.982	.178	.052	.571	-.156	1.175	1.185	.131	.097
		-.871	.114	1.004	0.982	.129	.051	.541	-.119	1.127	1.151	.114	.089
		-.792	.067	0.981	0.967	.112	.045	.528	-.080	1.123	1.153	.114	.086
	100	-.791	.268	1.060	0.987	.134	.058	.149	-.091	1.136	1.055	.088	.065
		-.704	.246	0.914	0.959	.067	.050	.177	-.060	1.128	1.057	.087	.065
		-.654	.222	0.847	0.965	.046	.048	.156	-.079	1.109	1.045	.080	.064
	200	-.584	.071	1.023	1.046	.096	.060	.468	-.075	1.030	1.089	.081	.072
		-.591	.029	0.989	1.043	.086	.062	.474	-.043	1.022	1.079	.081	.069
		-.554	.002	0.981	1.090	.074	.071	.450	-.024	.976	1.031	.068	.056
	500	-.393	.063	1.010	1.029	.068	.058	.265	-.068	1.013	1.034	.060	.058
		-.383	.063	1.008	1.042	.072	.061	.260	-.067	1.016	1.036	.064	.061
		-.359	.057	1.006	1.066	.067	.069	.265	-.046	1.001	1.022	.058	.056
-.5	50	-.818	.220	1.180	1.053	.157	.071	.281	-.165	1.182	1.124	.102	.083
		-.668	.211	0.979	0.974	.083	.056	.293	-.126	1.155	1.119	.101	.080
		-.602	.185	0.915	0.927	.066	.046	.275	-.136	1.141	1.117	.095	.078
	100	-.592	.200	1.060	1.043	.098	.066	.109	-.061	1.049	1.055	.063	.065
		-.497	.187	0.936	1.032	.051	.061	.097	-.065	1.031	1.037	.057	.058
		-.453	.139	0.908	1.079	.047	.071	.109	-.043	1.040	1.046	.060	.061
	200	-.660	.205	1.033	1.028	.113	.062	.154	-.076	1.071	1.022	.070	.057
		-.578	.211	0.927	0.996	.059	.054	.120	-.104	1.082	1.034	.074	.059
		-.551	.143	0.868	1.063	.044	.068	.131	-.077	1.049	1.014	.063	.055
	500	-.350	.067	1.009	1.033	.067	.059	.267	-.059	1.004	1.031	.060	.057
		-.335	.069	0.995	1.040	.063	.063	.259	-.060	1.010	1.036	.064	.059
		-.342	.022	0.963	1.062	.050	.064	.264	-.031	0.996	1.022	.055	.057

Table 11. Performance of t-ratios: Group Interaction with $k = n^{0.5}$, `MRSAR-C`, $\sigma = 2$
Three rows under each n: `dgp1-dgp3`

		Each two columns for: $t(\lambda)$ and $t^{\text{bc}}(\lambda)$						Each two columns for: $t(\beta_1)$ and $t^{\text{bc}}(\beta_1)$					
λ	n	mean		sd		tail prob		mean		sd		tail prob	
0.5	50	-.603	.100	1.050	0.996	.106	.052	.396	-.038	1.102	1.086	.094	.072
		-.604	.052	0.998	1.000	.095	.050	.392	-.009	1.075	1.064	.084	.068
		-.553	.038	0.952	0.992	.072	.051	.338	-.029	1.029	1.028	.071	.060
	100	-.551	.115	1.034	1.003	.092	.054	.351	-.101	1.059	1.050	.078	.064
		-.539	.096	0.971	0.991	.075	.052	.355	-.073	1.054	1.047	.078	.064
		-.507	.069	0.917	1.007	.057	.056	.336	-.057	1.010	1.010	.068	.054
	200	-.378	.089	1.016	1.019	.072	.056	.274	-.065	1.031	1.040	.067	.063
		-.378	.075	0.981	1.021	.064	.057	.271	-.054	1.011	1.020	.060	.056
		-.361	.061	0.940	1.048	.051	.061	.255	-.050	0.979	0.991	.052	.048
	500	-.313	.074	1.008	1.015	.064	.056	.209	-.070	1.015	1.022	.062	.057
		-.311	.071	0.985	1.015	.057	.054	.211	-.061	1.016	1.022	.059	.054
		-.322	.037	0.936	1.025	.049	.059	.215	-.040	0.982	0.987	.053	.046
0.2	50	-.676	.121	1.079	0.974	.125	.047	.314	-.080	1.154	1.104	.101	.078
		-.641	.104	1.015	0.977	.102	.049	.311	-.066	1.129	1.092	.096	.076
		-.610	.081	0.971	0.972	.081	.048	.288	-.076	1.098	1.074	.086	.072
	100	-.531	.119	1.024	0.986	.087	.050	.321	-.089	1.055	1.050	.072	.063
		-.526	.102	0.997	1.007	.080	.053	.312	-.080	1.060	1.057	.074	.065
		-.511	.073	0.972	1.049	.070	.063	.293	-.081	1.020	1.028	.065	.059
	200	-.442	.084	0.998	0.988	.071	.050	.316	-.066	1.023	1.027	.066	.057
		-.453	.060	0.992	1.008	.072	.054	.317	-.052	1.022	1.026	.067	.057
		-.429	.057	0.975	1.035	.063	.060	.296	-.052	0.999	1.006	.060	.054
	500	-.332	.054	1.018	1.022	.067	.060	.254	-.034	1.013	1.019	.062	.055
		-.322	.059	1.002	1.021	.061	.055	.247	-.036	1.009	1.016	.058	.054
		-.325	.042	0.987	1.047	.060	.064	.243	-.028	0.996	1.004	.056	.054
-.2	50	-.674	.026	1.171	1.177	.128	.079	.118	-.060	1.151	1.097	.089	.072
		-.547	.071	1.018	1.114	.077	.069	.095	-.081	1.143	1.096	.088	.075
		-.470	.111	0.938	1.000	.054	.050	.072	-.110	1.134	1.093	.086	.076
	100	-.527	.118	1.048	0.996	.093	.053	.258	-.066	1.073	1.055	.076	.064
		-.500	.087	0.972	0.994	.072	.050	.235	-.072	1.063	1.049	.073	.064
		-.441	.086	0.947	1.031	.055	.058	.235	-.057	1.020	1.015	.060	.054
	200	-.430	.108	1.024	1.004	.075	.054	.172	-.084	1.033	1.016	.060	.056
		-.420	.087	0.981	1.002	.067	.054	.189	-.061	1.036	1.021	.063	.054
		-.381	.076	0.928	1.023	.045	.056	.174	-.060	1.011	1.000	.058	.051
	500	-.319	.065	0.998	1.003	.059	.051	.243	-.061	1.013	1.022	.062	.056
		-.303	.069	0.984	1.007	.058	.052	.236	-.064	1.000	1.011	.057	.054
		-.283	.059	0.959	1.035	.048	.060	.211	-.070	0.985	0.998	.053	.051
-.5	50	-.685	.154	1.163	1.032	.133	.058	.136	-.071	1.167	1.094	.096	.075
		-.529	.194	0.986	0.955	.070	.045	.133	-.075	1.169	1.105	.097	.074
		-.427	.243	0.917	0.887	.050	.043	.108	-.115	1.151	1.096	.090	.076
	100	-.545	.117	1.061	0.997	.098	.050	.230	-.081	1.084	1.054	.078	.064
		-.435	.133	0.933	0.975	.052	.048	.196	-.099	1.059	1.035	.066	.058
		-.359	.137	0.862	0.963	.034	.047	.188	-.094	1.029	1.012	.059	.053
	200	-.453	.093	1.026	1.006	.077	.053	.273	-.070	1.041	1.041	.066	.060
		-.414	.069	0.949	1.004	.054	.054	.268	-.057	1.014	1.018	.061	.053
		-.329	.091	0.876	1.014	.035	.056	.217	-.081	0.965	0.973	.047	.046
	500	-.320	.062	1.012	1.014	.064	.052	.245	-.056	1.020	1.030	.062	.056
		-.293	.066	0.982	1.019	.055	.056	.234	-.058	1.004	1.015	.056	.054
		-.262	.044	0.908	1.023	.035	.060	.210	-.057	0.948	0.961	.042	.041

Table 12. Empirical Means and rmses of the MLE and Bias-Corrected MLEs of λ
Replication of the Monte Carlo Experiment of Bao and Ullah (2007a, p. 405)

n	J	λ	$\hat{\lambda}_n$	rmse	$\hat{\lambda}_n^{\text{ba2}}$	rmse	$\hat{\lambda}_n^{\text{bc2}}$	rmse	$\hat{\lambda}_n^{\text{bu2}}$	rmse
30	2	0.9	0.880	0.060	0.887	0.059	0.887	0.059	0.898	0.054
		0.4	0.381	0.159	0.383	0.160	0.383	0.160	0.398	0.164
		0.2	0.190	0.177	0.191	0.179	0.190	0.179	0.199	0.185
		0.0	0.003	0.180	0.003	0.181	0.002	0.181	0.003	0.189
		-0.2	-0.189	0.173	-0.190	0.174	-0.191	0.174	-0.198	0.180
		-0.4	-0.377	0.163	-0.380	0.164	-0.380	0.164	-0.395	0.168
		-0.9	-0.881	0.060	-0.888	0.059	-0.888	0.059	-0.899	0.054
	6	0.9	0.854	0.117	0.886	0.100	0.886	0.100	0.897	0.095
		0.4	0.318	0.280	0.379	0.265	0.377	0.265	0.393	0.272
		0.2	0.127	0.309	0.190	0.300	0.189	0.301	0.199	0.312
		0.0	-0.070	0.329	-0.007	0.325	-0.008	0.326	-0.006	0.340
		-0.2	-0.255	0.347	-0.193	0.350	-0.195	0.349	-0.201	0.367
		-0.4	-0.443	0.353	-0.386	0.362	-0.388	0.361	-0.404	0.380
		-0.9	-0.899	0.339	-0.858	0.357	-0.860	0.357	-0.900	0.370
	10	0.9	0.821	0.191	0.880	0.153	0.879	0.153	0.893	0.148
		0.4	0.258	0.389	0.375	0.354	0.372	0.354	0.390	0.365
		0.2	0.057	0.432	0.180	0.407	0.177	0.408	0.189	0.425
		0.0	-0.143	0.461	-0.018	0.448	-0.021	0.449	-0.018	0.471
		-0.2	-0.325	0.480	-0.203	0.481	-0.206	0.482	-0.212	0.508
		-0.4	-0.504	0.489	-0.388	0.504	-0.391	0.505	-0.408	0.533
		-0.9	-0.949	0.491	-0.862	0.534	-0.865	0.534	-0.908	0.561
100	2	0.9	0.895	0.027	0.898	0.026	0.898	0.026	0.900	0.026
		0.4	0.394	0.087	0.396	0.087	0.396	0.087	0.400	0.088
		0.2	0.198	0.097	0.199	0.098	0.199	0.097	0.201	0.098
		0.0	-0.001	0.100	-0.001	0.101	-0.001	0.101	-0.001	0.102
		-0.2	-0.197	0.097	-0.198	0.098	-0.198	0.098	-0.200	0.099
		-0.4	-0.395	0.087	-0.397	0.087	-0.397	0.087	-0.401	0.087
		-0.9	-0.895	0.027	-0.898	0.026	-0.898	0.026	-0.900	0.026
	6	0.9	0.887	0.043	0.897	0.040	0.897	0.040	0.899	0.039
		0.4	0.374	0.137	0.394	0.134	0.394	0.134	0.398	0.134
		0.2	0.175	0.160	0.196	0.158	0.196	0.158	0.199	0.159
		0.0	-0.023	0.177	-0.002	0.177	-0.002	0.177	-0.002	0.178
		-0.2	-0.216	0.186	-0.196	0.187	-0.195	0.187	-0.198	0.189
		-0.4	-0.416	0.196	-0.398	0.198	-0.398	0.198	-0.403	0.201
		-0.9	-0.902	0.195	-0.891	0.199	-0.891	0.199	-0.902	0.201
	10	0.9	0.880	0.058	0.897	0.051	0.897	0.051	0.899	0.050
		0.4	0.360	0.177	0.398	0.169	0.399	0.169	0.402	0.170
		0.2	0.154	0.213	0.195	0.207	0.195	0.207	0.198	0.208
		0.0	-0.046	0.236	-0.004	0.232	-0.004	0.232	-0.004	0.234
		-0.2	-0.243	0.259	-0.201	0.258	-0.202	0.258	-0.204	0.261
		-0.4	-0.430	0.267	-0.391	0.271	-0.390	0.271	-0.395	0.274
		-0.9	-0.915	0.288	-0.887	0.297	-0.887	0.297	-0.899	0.301

Note: $\hat{\lambda}_n^{\text{bu2}}$: the second-order analytically bias-corrected MLE of Bao and Ullah (2007a).
rmse: the empirical root mean squared error.
DGP: a pure SAR model, i.e., $Y_n = \lambda W_n + u_n,\ u_n \sim N(0, I_n)$.

Table 13. MLE and 2nd-Order Bias-Corrected MLE of SAR Models: **Replication I of Lee (2004a)**
Each Pair of Rows: Empirical Means and sds; Each Pair of Columns: MLE and Bias-Corrected MLE

R	θ	$m=3$		$m=5$		$m=10$		$m=20$		$m=50$		$m=100$	
SAR: $Y_n = \lambda W_n + u_n,\ u_n \sim N(0, \sigma^2 I_n),\ \lambda = 0.5$ and $\sigma = 1$													
30	λ	0.494	0.497	0.490	0.497	0.488	0.498	0.487	0.498	0.486	0.499	0.487	0.500
		0.067	0.067	0.068	0.067	0.068	0.066	0.068	0.067	0.070	0.068	0.068	0.067
	σ	0.992	0.991	0.995	0.994	0.998	0.997	0.999	0.998	1.000	0.999	1.000	0.999
		0.079	0.079	0.059	0.059	0.042	0.042	0.029	0.029	0.018	0.018	0.013	0.013
60	λ	0.496	0.498	0.496	0.499	0.494	0.499	0.494	0.500	0.494	0.500	0.493	0.499
		0.048	0.047	0.046	0.046	0.047	0.047	0.047	0.047	0.047	0.046	0.047	0.047
	σ	0.997	0.996	0.998	0.997	0.998	0.998	0.999	0.999	1.000	1.000	1.000	1.000
		0.056	0.056	0.042	0.042	0.029	0.029	0.020	0.020	0.013	0.013	0.009	0.009
120	λ	0.498	0.499	0.497	0.499	0.497	0.499	0.497	0.500	0.498	0.500	0.498	0.501
		0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.032
	σ	0.998	0.997	0.998	0.998	0.999	0.999	1.000	0.999	1.000	1.000	1.000	1.000
		0.040	0.040	0.030	0.030	0.021	0.021	0.015	0.015	0.009	0.009	0.007	0.007
MRSAR-1: $Y_n = \lambda W_n + X_n\beta + u_n$, where $u_n \sim N(0, \sigma^2 I_n)$, $X_n \sim N(0, I_n)$, $\lambda = .5, \beta = 1$, and $\sigma = 1$													
30	λ	0.496	0.497	0.492	0.498	0.492	0.500	0.491	0.499	0.490	0.499	0.490	0.499
		0.056	0.056	0.056	0.055	0.057	0.055	0.057	0.056	0.058	0.056	0.058	0.056
	β	0.997	0.996	0.999	0.998	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000
		0.108	0.108	0.082	0.082	0.058	0.058	0.041	0.041	0.026	0.026	0.018	0.018
	σ	0.986	0.986	0.992	0.991	0.996	0.995	0.998	0.997	0.999	0.999	1.000	0.999
		0.080	0.080	0.059	0.059	0.042	0.042	0.029	0.029	0.018	0.018	0.013	0.013
60	λ	0.497	0.497	0.497	0.499	0.496	0.500	0.495	0.499	0.496	0.500	0.496	0.500
		0.039	0.039	0.038	0.038	0.038	0.038	0.039	0.038	0.038	0.038	0.039	0.038
	β	1.000	0.999	1.000	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
		0.077	0.077	0.059	0.059	0.041	0.041	0.029	0.029	0.019	0.019	0.013	0.013
	σ	0.994	0.993	0.996	0.996	0.998	0.997	0.999	0.999	0.999	0.999	1.000	1.000
		0.055	0.055	0.042	0.042	0.029	0.029	0.020	0.020	0.013	0.013	0.009	0.009
120	λ	0.499	0.499	0.498	0.499	0.499	0.500	0.498	0.500	0.498	0.500	0.498	0.500
		0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027
	β	1.000	1.000	1.000	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
		0.054	0.054	0.041	0.041	0.029	0.029	0.021	0.021	0.013	0.013	0.009	0.009
	σ	0.996	0.996	0.998	0.998	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000
		0.039	0.039	0.030	0.030	0.021	0.021	0.015	0.015	0.009	0.009	0.006	0.006
MRSAR-2: $Y_n = \lambda W_n + X_n\beta + u_n$, $u_n \sim N(0, \sigma^2 I_n)$, the elements x_{ir} of X_n are $x_{ir} = (z_r + z_{ir})/\sqrt{2}$ with $z_r's$ and $z_{ir}'s$ being iid $N(0,1)$, $\lambda = .5, \beta = 1$, and $\sigma = 1$.													
30	λ	0.494	0.496	0.493	0.499	0.493	0.500	0.493	0.499	0.496	0.500	0.498	0.500
		0.059	0.058	0.055	0.055	0.052	0.051	0.045	0.044	0.034	0.034	0.027	0.027
	β	1.003	1.001	1.001	0.998	1.001	0.998	1.001	0.999	1.001	1.000	1.000	0.999
		0.138	0.138	0.105	0.105	0.075	0.075	0.053	0.053	0.034	0.034	0.024	0.024
	σ	0.986	0.985	0.992	0.991	0.996	0.996	0.998	0.998	0.999	0.999	1.000	0.999
		0.079	0.079	0.060	0.060	0.041	0.041	0.029	0.029	0.018	0.018	0.013	0.013
60	λ	0.497	0.498	0.496	0.499	0.496	0.500	0.497	0.500	0.498	0.500	0.499	0.500
		0.041	0.041	0.039	0.039	0.036	0.035	0.031	0.031	0.024	0.024	0.019	0.019
	β	1.000	1.000	1.002	1.000	1.000	0.999	1.001	0.999	1.000	0.999	1.000	1.000
		0.096	0.096	0.074	0.074	0.053	0.053	0.038	0.038	0.024	0.024	0.018	0.018
	σ	0.993	0.993	0.996	0.995	0.999	0.998	0.999	0.999	0.999	0.999	1.000	1.000
		0.056	0.056	0.042	0.042	0.029	0.029	0.021	0.021	0.013	0.013	0.009	0.009
120	λ	0.499	0.499	0.498	0.499	0.498	0.500	0.499	0.500	0.499	0.500	0.499	0.500
		0.029	0.029	0.027	0.027	0.025	0.025	0.022	0.022	0.017	0.017	0.013	0.013
	β	1.000	1.000	1.000	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
		0.068	0.068	0.052	0.053	0.037	0.037	0.027	0.027	0.017	0.017	0.012	0.012
	σ	0.996	0.996	0.998	0.998	0.999	0.999	1.000	0.999	1.000	1.000	1.000	1.000
		0.039	0.039	0.030	0.030	0.020	0.020	0.015	0.015	0.009	0.009	0.007	0.007

Table 14. MLE and 2nd-Order Bias-Corrected MLE of SAR Models: **Replication II of Lee (2004a)**
Each Pair of Rows: Empirical Means and sds; Each Pair of Columns: MLE and Bias-Corrected MLE

m	θ	$R = 3$		$R = 5$		$R = 10$		$R = 20$		$R = 50$		$R = 100$	
SAR: $Y_n = \lambda W_n + u_n,\ u_n \sim N(0, \sigma^2 I_n),\ \lambda = 0.5$ and $\sigma = 1$													
30	λ	0.327	0.478	0.409	0.490	0.460	0.498	0.481	0.499	0.492	0.499	0.497	0.500
		0.429	0.331	0.238	0.206	0.136	0.127	0.087	0.084	0.052	0.051	0.036	0.036
	σ	0.991	0.990	0.996	0.994	0.997	0.996	0.999	0.998	1.000	0.999	1.000	0.999
		0.076	0.075	0.058	0.058	0.042	0.042	0.029	0.029	0.019	0.019	0.013	0.013
60	λ	0.323	0.484	0.406	0.491	0.461	0.500	0.480	0.498	0.493	0.500	0.496	0.500
		0.441	0.337	0.249	0.213	0.136	0.126	0.088	0.085	0.052	0.051	0.035	0.035
	σ	0.996	0.995	0.997	0.997	0.999	0.998	0.999	0.999	1.000	1.000	1.000	1.000
		0.053	0.053	0.041	0.041	0.029	0.029	0.021	0.021	0.013	0.013	0.009	0.009
120	λ	0.307	0.476	0.405	0.492	0.457	0.497	0.480	0.499	0.493	0.500	0.497	0.500
		0.514	0.375	0.253	0.216	0.137	0.127	0.087	0.084	0.052	0.051	0.035	0.035
	σ	0.998	0.997	0.998	0.998	0.999	0.999	1.000	0.999	1.000	1.000	1.000	1.000
		0.037	0.037	0.029	0.029	0.021	0.021	0.014	0.014	0.009	0.009	0.006	0.006
MRSAR-1: $Y_n = \lambda W_n + X_n\beta + u_n$, where $u_n \sim N(0, \sigma^2 I_n)$, $X_n \sim N(0, I_n)$, $\lambda = .5, \beta = 1$, and $\sigma = 1$													
30	λ	0.365	0.473	0.435	0.493	0.471	0.499	0.486	0.500	0.494	0.499	0.497	0.500
		0.369	0.282	0.200	0.169	0.112	0.103	0.071	0.068	0.043	0.042	0.029	0.029
	β	0.998	0.996	0.999	0.997	1.000	0.999	1.000	0.999	1.000	1.000	1.000	1.000
		0.107	0.107	0.083	0.083	0.058	0.058	0.041	0.041	0.026	0.026	0.019	0.019
	σ	0.984	0.983	0.991	0.990	0.996	0.995	0.998	0.998	0.999	0.999	1.000	1.000
		0.076	0.076	0.059	0.059	0.041	0.041	0.029	0.029	0.018	0.018	0.013	0.013
60	λ	0.365	0.481	0.432	0.494	0.470	0.498	0.487	0.500	0.494	0.500	0.497	0.500
		0.381	0.289	0.199	0.166	0.112	0.102	0.072	0.069	0.043	0.042	0.030	0.029
	β	0.998	0.997	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.075	0.075	0.057	0.057	0.041	0.041	0.029	0.029	0.018	0.018	0.013	0.013
	σ	0.993	0.992	0.996	0.995	0.998	0.998	0.999	0.999	1.000	1.000	1.000	1.000
		0.054	0.054	0.041	0.041	0.029	0.029	0.021	0.021	0.013	0.013	0.009	0.009
120	λ	0.369	0.488	0.434	0.497	0.468	0.498	0.486	0.500	0.494	0.500	0.497	0.500
		0.368	0.275	0.201	0.168	0.114	0.104	0.071	0.068	0.042	0.041	0.029	0.029
	β	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.053	0.053	0.041	0.041	0.029	0.029	0.020	0.020	0.013	0.013	0.009	0.009
	σ	0.997	0.996	0.998	0.998	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000
		0.037	0.037	0.029	0.029	0.021	0.021	0.014	0.014	0.009	0.009	0.007	0.007
MRSAR-2: $Y_n = \lambda W_n + X_n\beta + u_n$, $u_n \sim N(0, \sigma^2 I_n)$, the elements x_{ir} of X_n are $x_{ir} = (z_r + z_{ir})/\sqrt{2}$ with $z_r{}'s$ and $z_{ir}{}'s$ being iid $N(0,1)$, $\lambda = .5, \beta = 1$, and $\sigma = 1$.													
30	λ	0.428	0.489	0.462	0.495	0.482	0.498	0.492	0.500	0.497	0.500	0.499	0.500
		0.221	0.182	0.131	0.118	0.078	0.074	0.051	0.050	0.031	0.030	0.022	0.021
	β	1.010	0.992	1.007	0.996	1.002	0.996	1.001	0.998	1.000	0.999	1.000	1.000
		0.143	0.144	0.109	0.110	0.076	0.076	0.054	0.054	0.034	0.034	0.024	0.024
	σ	0.986	0.986	0.992	0.991	0.996	0.996	0.998	0.998	0.999	0.999	0.999	0.999
		0.077	0.077	0.059	0.059	0.041	0.041	0.029	0.029	0.018	0.018	0.013	0.013
60	λ	0.447	0.490	0.475	0.497	0.488	0.499	0.494	0.499	0.498	0.500	0.499	0.500
		0.186	0.156	0.096	0.088	0.061	0.059	0.040	0.040	0.025	0.025	0.017	0.017
	β	1.007	0.995	1.004	0.996	1.002	0.998	1.002	1.000	1.000	1.000	1.000	1.000
		0.102	0.103	0.077	0.078	0.055	0.055	0.038	0.039	0.025	0.025	0.017	0.017
	σ	0.993	0.993	0.996	0.996	0.998	0.998	0.999	0.999	1.000	1.000	1.000	1.000
		0.053	0.053	0.041	0.041	0.028	0.028	0.020	0.020	0.013	0.013	0.009	0.009
120	λ	0.469	0.496	0.487	0.501	0.493	0.499	0.497	0.500	0.499	0.500	0.499	0.500
		0.137	0.118	0.071	0.068	0.045	0.044	0.030	0.030	0.019	0.019	0.013	0.013
	β	1.004	0.997	1.002	0.997	1.001	0.998	1.001	0.999	1.000	1.000	1.000	1.000
		0.072	0.072	0.055	0.055	0.038	0.038	0.028	0.028	0.017	0.017	0.012	0.012
	σ	0.996	0.996	0.998	0.998	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000
		0.038	0.038	0.029	0.029	0.020	0.020	0.014	0.014	0.009	0.009	0.007	0.007