# Improved Inferences for Spatial Regression Models*

Shew Fan Liu† and Zhenlin Yang

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

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

August 25, 2015

**Abstract**

The quasi-maximum likelihood (QML) method is popular in the estimation and inference for spatial regression models. However, the QML estimators (QMLEs) of the spatial parameters can be quite biased and hence the standard inferences for the regression coefficients (based on $t$-ratios) can be seriously affected. This issue, however, has not been addressed. The QMLEs of the spatial parameters can be bias-corrected based on the general method of Yang (2015b, *J. of Econometrics* 186, 178-200). In this paper, we demonstrate that by simply replacing the QMLEs of the spatial parameters by their bias-corrected versions, the usual $t$-ratios for the regression coefficients can be greatly improved. We propose further corrections on the standard errors of the QMLEs of the regression coefficients, and the resulted $t$-ratios perform superbly, leading to much more reliable inferences.

**Key Words:** Asymptotic inference; Bias correction; Bootstrap; Improved $t$-ratio; Monte Carlo; Spatial layout; Stochastic expansion; Variance correction.

**JEL Classification:** C10, C12, C15, C21

## 1. Introduction

The maximum likelihood (ML) or quasi-ML (QML) method is popular in the estimation and inference for spatial regression models (Anselin, 1988; Anselin and Bera, 1998; Lee, 2004). However, the ML estimators (MLEs) or quasi-MLEs (QMLEs) of the spatial parameters can be quite biased (Bao and Ullah, 2007; Yang, 2015b; Liu and Yang, 2015) and hence the standard inferences for spatial effects and covariate effects, based on LM-statistics or $t$-statistics referring to the asymptotic standard normal distribution, can be seriously affected. Much effort has been devoted recently to the development of improved inference methods for the spatial econometrics models. However, most of the research has been focused on improving inferences for spatial effects in the form of point estimation (Bao and Ullah, 2007; Bao, 2013; Liu and Yang, 2015; Yang, 2015b) and testing (Baltagi and Yang, 2013a,b; Robinson and Rossi, 2014a,b; Yang, 2010;

*We thank the Editor Daniel McMillen and two referees for the helpful comments and suggestions that improved the paper. Thanks are also due to the participants of the *14th International Workshop on Spatial Econometrics and Statistics*, Paris, 27-28 May 2015, for their useful comments. Zhenlin Yang gratefully acknowledges the support from a research grant (C244/MSS14E002) from Singapore Management University.

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

Yang, 2015a,b). Little or no attention has been paid to the development of improved inferences for the covariate effects in the spatial regression models.

Yang (2015a) proposed a general method for constructing 2nd-order accurate bootstrap LM tests for spatial effects, but the issue of improved inferences for covariate effects was not studied. Yang (2015b) proposed a general method for 3rd-order bias and variance corrections on nonlinear estimators which are prone to finite sample bias, and argued that once the biases of nonlinear estimators are corrected, the biases of covariate effects and error standard deviations become negligible. He demonstrated the effectiveness of the methods using the linear regression model with spatial lag dependence with results showing that a 2nd-order bias correction is largely sufficient. He further demonstrated that the 2nd-order or 3rd-order corrected $t$-statistics for spatial effect indeed improve upon the standard $t$-statistics greatly, but again, no study was carried out to test the performance of the $t$-statistics for covariate effects, and its improvements.

Evidently, in practical applications of spatial econometrics models, it is central to have a set of reliable inference methods for the covariate effects. In this paper, we adopt the bias-correction method of Yang (2015b) to propose methods that 'correct' the standard $t$-statistics for the regression coefficients. We demonstrate that by simply replacing the QMLEs of the spatial parameters by their bias-corrected versions, the usual $t$-ratios for the regression coefficients can be greatly improved. We propose further corrections on the standard errors of the 'bias-corrected' QMLEs of the regression coefficients, and the resulted $t$-ratios perform superbly, leading to much more reliable inferences. The proposed methods are simple and can be easily adopted by practitioners. We consider in detail three popular spatial regression models: the linear regression model with spatial error dependence (SED), that with a spatial lag dependence (SLD), and that with both SLD and SED, also referred to as the SARAR model in the literature. See Anselin and Bera (1998) and Anselin (2001) for excellent reviews on these models. Bias-correction on a single spatial estimator has been considered in detail in Yang (2015b) for the SLD model, and in Liu and Yang (2015b) for the SED model. Bias-corrections for the SARAR model have not been formally considered, although briefly discussed in Yang (2015b) under a general outline for bias corrections for a model with a vector of non-linear parameters.

The line-up for the paper is as follows. Section 2 outlines the general method of bias correction on nonlinear estimators, and the methods for constructing improved $t$-statistics for the linear parameters in the model. Sections 3-5 study in detail the improved inference methods for the regression coefficients for, respectively, the SED model, the SLD model, and the SARAR model. Each of Sections 3-5 is accompanied with a set of Monte Carlo simulation results. Section 6 concludes the paper, and discuss further extensions of the proposed methodology.

## 2. Method of Bias Correction for Nonlinear Estimation

From the discussions in the introduction, it is clear that the key for an improved inference for the regression coefficients is to bias-correct the QMLEs of the spatial parameters in a spatial regression model. We now outline the method of bias correction on nonlinear estimators, not necessarily the QMLEs of the spatial parameters. In studying the finite sample properties of a parameter estimator, say $\hat{\theta}_n$, defined as $\hat{\theta}_n = \arg\{\psi_n(\theta) = 0\}$ for a *joint estimating function* (JEF) $\psi_n(\theta)$, based on a sample of size $n$, Rilstone et al. (1996) developed a stochastic expansion from which a bias-correction on $\hat{\theta}_n$ can be made. The vector of parameters $\theta$ may contain a set of *linear and scale parameters*, say $\alpha$, and a few *non-linear parameters*, say $\delta$, in the sense that given $\delta$, the constrained estimator $\tilde{\alpha}_n(\delta)$ of the vector $\alpha$ possesses an explicit expression but the estimation of $\delta$ has to be done through numerical optimization. In this case, Yang (2015b) argued that it is more effective to work with the *concentrated estimating function* (CEF): $\tilde{\psi}_n(\delta) = \psi_n(\tilde{\alpha}_n(\delta), \delta)$, and to perform a stochastic expansion based on this CEF and hence bias corrections on the non-linear estimators defined by,

$$\hat{\delta}_n = \arg\{\tilde{\psi}_n(\delta) = 0\}, \tag{1}$$

which not only reduces the dimensionality of the bias-correction problem (a multi-dimensional problem is reduced to a single-dimensional problem if $\delta$ is a scalar parameter), but also takes into account the additional variability from the estimation of the 'nuisance' parameters $\alpha$.

Let $H_{rn}(\delta) = \nabla^r \tilde{\psi}_n(\delta), r = 1, 2, 3$, be the partial derivatives of $\tilde{\psi}_n(\delta)$, carried out sequentially and elementwise with respect to $\delta'$, $\tilde{\psi}_n \equiv \tilde{\psi}_n(\delta_0)$, $H_{rn} \equiv H_{rn}(\delta_0)$, $H^{\circ}_{rn} = H_{rn} - \mathrm{E}(H_{rn}), r = 1, 2, 3$, and $\Omega_n = -[\mathrm{E}(H_{1n})]^{-1}$. Yang (2015b) presents a set of sufficient conditions under which $\hat{\delta}_n$ possesses the following third-order stochastic expansion at $\delta_0$, the true value of $\delta$:

$$\hat{\delta}_n - \delta_0 = a_{-1/2} + a_{-1} + a_{-3/2} + O_p(n^{-2}), \tag{2}$$

where, $a_{-s/2}$ represents terms of order $O_p(n^{-s/2})$ for $s = 1, 2, 3$, having the expressions,

$$\begin{aligned}
a_{-1/2} &= \Omega_n \tilde{\psi}_n, \\
a_{-1} &= \Omega_n H^{\circ}_{1n} a_{-1/2} + \tfrac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2} \otimes a_{-1/2}), \\
a_{-3/2} &= \Omega_n H^{\circ}_{1n} a_{-1} + \tfrac{1}{2}\Omega_n H^{\circ}_{2n}(a_{-1/2} \otimes a_{-1/2}) \\
&\quad + \tfrac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2} \otimes a_{-1} + a_{-1} \otimes a_{-1/2}) \\
&\quad + \tfrac{1}{6}\Omega_n \mathrm{E}(H_{3n})(a_{-1/2} \otimes a_{-1/2} \otimes a_{-1/2}),
\end{aligned}$$

with $\otimes$ denoting the Kronecker product.

When $\delta$ is a scalar, $a_{-s/2}$ simplifies to: $a_{-1/2} = \Omega_n \tilde{\psi}_n$, $a_{-1} = \Omega_n H^{\circ}_{1n} a_{-1/2} + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a^2_{-1/2})$, and $a_{-3/2} = \Omega_n H^{\circ}_{1n} a_{-1} + \frac{1}{2}\Omega_n H^{\circ}_{2n}(a^2_{-1/2}) + \Omega_n \mathrm{E}(H_{2n})(a_{-1/2} a_{-1}) + \frac{1}{6}\Omega_n \mathrm{E}(H_{3n})(a^3_{-1/2})$.

The key difference between the CEF-based and JEF-based expansions is that $\mathrm{E}[\tilde{\psi}_n(\delta_0)] \neq 0$ in general, but $\mathrm{E}[\psi_n(\theta_0)] = 0$, which allows a CEF-based bias correction to be derived under a more relaxed condition. Thus, a third-order expansion for the bias of $\hat{\delta}_n$ takes the form:

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

where $b_{-1} = \mathrm{E}(a_{-1/2} + a_{-1})$ and $b_{-3/2} = \mathrm{E}(a_{-3/2})$, being respectively the second- and third-order biases of $\hat{\delta}_n$. If an estimator $\hat{b}_{-1}$ of $b_{-1}$ is available such that $\mathrm{Bias}(\hat{b}_{-1}) = O(n^{-3/2})$, then a second-order bias-corrected estimator of $\delta$ is,

$$\delta_n^{\mathrm{bc2}} = \hat{\delta}_n - \hat{b}_{-1}. \tag{4}$$

If estimators $\hat{b}_{-1}$ and $\hat{b}_{-3/2}$ of both $b_{-1}$ and $b_{-3/2}$ are available such that $\mathrm{Bias}(\hat{b}_{-1}) = O(n^{-2})$ and $\mathrm{Bias}(\hat{b}_{-3/2}) = O(n^{-2})$, we have a third-order bias-corrected estimator of $\delta$ as,

$$\delta_n^{\mathrm{bc3}} = \hat{\delta}_n - \hat{b}_{-1} - \hat{b}_{-3/2}. \tag{5}$$

An obvious approach for finding the feasible corrections $\hat{b}_{-1}$ and $\hat{b}_{-3/2}$ is to first find the analytical expressions for $b_{-1}$ and $b_{-3/2}$ and then plugging in $\hat{\theta}_n$ for $\theta_0$. This approach is generally not feasible for two reasons: first, it is often difficult to find these analytical expressions even for known error distributions, and second, even if these expressions are available, it may involve higher-order moments of the errors if they are nonnormal, for which estimation may be unstable numerically. To overcome this difficulty, Yang (2015b) proposed a simple and yet very effective bootstrap method to estimate the relevant expected values.

Suppose that the model under consideration takes the form

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

and that the key quantities $\tilde{\psi}_n$ and $H_{rn}$ can be expressed as $\tilde{\psi}_n \equiv \tilde{\psi}_n(e_n, \theta_0)$ and $H_{rn} \equiv H_{rn}(e_n, \theta_0), r = 1, 2, 3$. Let $\hat{e}_n = g(Z_n, \hat{\theta}_n)$ be the vector of estimated residuals based on the original data, and $\hat{\mathcal{F}}_n$ be the empirical distribution function (EDF) of $\hat{e}_n$ (centered). When $\delta$ is a scalar parameter, the bootstrap estimates of the quantities in the bias terms are:

$$\widehat{\mathrm{E}}(\tilde{\psi}_n^i H_{rn}^j) = \mathrm{E}^*[\tilde{\psi}_n^i(\hat{e}_n^*, \hat{\theta}_n) H_{rn}^j(\hat{e}_n^*, \hat{\theta}_n)], \quad i, j = 0, 1, 2, \ldots, \quad r = 1, 2, 3, \tag{6}$$

where $\mathrm{E}^*$ denotes the expectation with respect to $\hat{\mathcal{F}}_n$, and $\hat{e}_n^*$ is a vector of $n$ random draws from $\hat{\mathcal{F}}_n$. To make (6) practically feasible, the following procedure can be followed.

**Bootstrap Algorithm 1 (BA-1):**

1. *Compute $\hat{\theta}_n$ defined by JEF, $\hat{e}_n = g(Z_n, \hat{\theta}_n)$, and EDF $\hat{\mathcal{F}}_n$ of the centered $\hat{e}_n$;*
2. *Draw a random sample of size $n$ from $\hat{\mathcal{F}}_n$, and denote the resampled vector by $\hat{e}_{n,b}^*$,*
3. *Compute $\tilde{\psi}_n(\hat{e}_{n,b}^*, \hat{\theta}_n)$ and $H_{rn}(\hat{e}_{n,b}^*, \hat{\theta}_n), r = 1, 2, 3;$*
4. *Repeat steps 2.-3. for $B$ times, to give approximate bootstrap estimates as,*

$$\mathrm{E}^*[\tilde{\psi}_n^i(\hat{e}_n^*, \hat{\theta}_n) H_{rn}^j(\hat{e}_n^*, \hat{\theta}_n)] \doteq \tfrac{1}{B} \textstyle\sum_{b=1}^{B} \tilde{\psi}_n^i(\hat{e}_{n,b}^*, \hat{\theta}_n) H_{rn}^j(\hat{e}_{n,b}^*, \hat{\theta}_n),$$

*for $i, j = 0, 1, 2, \ldots, \quad r = 1, 2, 3.$*

The approximations in the last step can be made arbitrarily accurate by choosing an arbitrarily large $B$. Yang (2015b) shows that under certain conditions:

$$\begin{aligned}\text{Bias}(\delta_n^{\text{bc2}}) &= \text{Bias}(\hat{\delta}_n) - \text{E}(\hat{b}_{-1}) \\ &= -\text{Bias}(\hat{b}_{-1}) + O(n^{-3/2}) = O(n^{-3/2}), \text{ and} \\ \text{Bias}(\delta_n^{\text{bc3}}) &= \text{Bias}(\hat{\delta}_n) - \text{E}(\hat{b}_{-1}) - \text{E}(\hat{b}_{-3/2}) \\ &= -\text{Bias}(\hat{b}_{-1}) - \text{Bias}(\hat{b}_{-3/2}) + O(n^{-2}) = O(n^{-2}).\end{aligned}$$

When $\delta$ becomes a vector, the non-stochastic and stochastic quantities are mixed in $b_{-1}$ and $b_{-3/2}$. In this case, Yang (2015b) proposed that instead of going through the algebraic procedure to separate the two types of quantities so that the expectations of various quantities can be bootstrapped in one round, the above bootstrap procedure can be revised as follows.

**Bootstrap Algorithm 2 (BA-2):**

1. *Draw $B$ independent random samples, $\{\hat{e}^*_{n,b},\ b = 1, 2, \ldots, B\}$, from $\hat{\mathcal{F}}_n$,*
2. *Calculate the bootstrap estimates of $E(H_{1n})$ and $E(H_{2n})$,*

$$\widehat{\text{E}}(H_{1n}) = \tfrac{1}{B}\textstyle\sum_{b=1}^n H_{1n}(\hat{e}^*_{n,b}, \hat{\theta}_n) \textit{ and } \widehat{\text{E}}(H_{2n}) = \tfrac{1}{B}\textstyle\sum_{b=1}^n H_{2n}(\hat{e}^*_{n,b}, \hat{\theta}_n)$$

3. *Based on the bootstrap estimates $\hat{\Omega}_n = -\widehat{\text{E}}^{-1}(H_{1n})$ and $\widehat{\text{E}}(H_{2n})$, calculate the bootstrap estimate of, e.g., $\text{E}[H^\circ_{2n}(a_{-1/2} \otimes a_{-1/2})]$, as*

$$\tfrac{1}{B}\textstyle\sum_{b=1}^n \left\{[H_{2n}(\hat{e}^*_{n,b}, \hat{\theta}_n) - \widehat{\text{E}}(H_{2n})][\hat{\Omega}_n\tilde{\psi}_n(\hat{e}^*_{n,b}, \hat{\theta}_n) \otimes \hat{\Omega}_n\tilde{\psi}_n(\hat{e}^*_{n,b}, \hat{\theta}_n)]\right\}.$$

The other quantities can be handled in a similar manner. This is essentially a two-round bootstrap procedure as it runs the iterations $b = 1, 2, \ldots, B$ two times, based on the same sequence of bootstrap samples. Computationally it is slightly more demanding, but algebraically it is much simpler and thus easier to code. As noted by Yang (2015b), these procedures are time-efficient as the reestimation of the parameters in the bootstrap process is avoided.

**Inferences following bias-correction.** There are mainly two types of inferences that could benefit from the bias-corrections on the nonlinear estimators: one is the inference for the nonlinear parameters, and the other for the linear parameters. In the framework of linear regressions with spatial dependence, the spatial parameters are the nonlinear parameters, and the regression coefficients are the linear parameters. Improved tests for spatial effects have been considered by Baltagi and Yang (2013a,b), Robinson and Rossi (2014a,b), Yang (2010), and Yang (2015a). However, the issue of improved inferences for the regression coefficients has not been considered, except that it is briefly mentioned in Liu and Yang (2015).

To fix the idea, we focus on the 2nd-order bias-corrected $\hat{\delta}_n$, the $\hat{\delta}_n^{\text{bc2}}$. Let $\hat{\alpha}_n \equiv \tilde{\alpha}_n(\hat{\delta}_n)$ and $\hat{\alpha}_n^{\text{bc}} \equiv \tilde{\alpha}_n(\hat{\delta}_n^{\text{bc2}})$, and $\hat{\theta}_n = (\hat{\alpha}_n', \hat{\delta}_n')'$ and $\hat{\theta}_n^{\text{bc}} = (\hat{\alpha}_n^{\text{bc}\prime}, \hat{\delta}_n^{\text{bc2}\prime})'$. Yang (2015b) argued that estimation of the nonlinear parameter is the main source of bias and once the nonlinear estimator is bias-corrected the resulting linear estimators would be nearly unbiased. Let $\Omega_n(\theta_0)$ be the asymptotic variance-covariance (VC) matrix of $\hat{\alpha}_n$. Then, an asymptotic $t$-statistic for inference for $c_0'\alpha_0$, a linear contrast of $\alpha_0$, has the familiar form:

$$t_n = (c_0'\hat{\alpha}_n - c_0'\alpha_0)/\sqrt{c_0'\Omega_n(\hat{\theta}_n)c_0}.$$

Simply replacing $\hat{\theta}_n$ by $\hat{\theta}_n^{\text{bc}}$, a possibly improved $t$-statistic results:

$$t_n^{\text{bc}} = (c_0'\hat{\alpha}_n^{\text{bc}} - c_0'\alpha_0)/\sqrt{c_0'\Omega_n(\hat{\theta}_n^{\text{bc}})c_0}.$$

The statistic $t_n^{\text{bc}}$ is not fully 2nd-order corrected as it uses the asymptotic variance of $\hat{\alpha}_n$ evaluated at $\hat{\theta}_n^{\text{bc}}$. Further, the estimator $\hat{\alpha}_n^{\text{bc}}$ is also not fully 2nd-order bias-corrected, although it can easily be made so. Let $\hat{\alpha}_n^{\text{bc2}}$ be the 2nd-order bias-corrected $\hat{\alpha}_n$ or $\hat{\alpha}_n^{\text{bc}}$. Let $\Omega_n^{\text{bc2}}(\theta_0)$ be the 2nd-order variance of $\hat{\alpha}_n^{\text{bc2}}$, and $\hat{\Omega}_n^{\text{bc2}}$ be its consistent estimate. A fully 2nd-order corrected $t$-statistic, using a 2nd-order bias-corrected estimator and its 2nd-order variance, is thus:

$$t_n^{\text{bc2}} = (c_0'\hat{\alpha}_n^{\text{bc2}} - c_0'\alpha_0)/\sqrt{c_0'\hat{\Omega}_n^{\text{bc2}}c_0}.$$

Typically, $\Omega_n^{\text{bc2}}(\theta_0)$ does not have an explicit expression, but the bootstrap methods described above can be extended to give a consistent estimate of it. See the subsequent sections for details.

# 3. Improved Inferences for the SED Model

In this section, we study the inference methods for the regression coefficients of the SED model. First, in Section 3.1, we outline the QML estimation for this model and inferences based on the asymptotic distribution of the QMLEs of the model parameters, then in Section 3.2 we outline the method of bias-correcting the QMLE of the spatial parameter, and then in Section 3.3 we present the improved inference methods. To assess the finite sample performance of the asymptotic and improved inferences, Monte Carlo results are presented in Section 3.4.

## 3.1 QML estimation and asymptotic inference

Consider the following linear regression model with spatial error dependence (SED), where the SED is specified as a spatial autoregressive (SAR) process:

$$Y_n = X_n\beta + u_n, \quad u_n = \rho W_n u_n + \epsilon_n, \tag{7}$$

where $Y_n$ is an $n \times 1$ vector of observations on the dependent variable corresponding to $n$ spatial units, $X_n$ is an $n \times k$ matrix containing the values of $k$ exogenous regressors, $W_n$ is an $n \times n$

spatial weight matrix that summarises the interactions among the spatial units, $\epsilon_n$ is an $n \times 1$ vector of independent and identically distributed (iid) disturbances with mean zero and variance $\sigma^2$, $\rho$ is the *spatial parameter*, and $\beta$ denotes the $k \times 1$ vector of regression coefficients.

The quasi Gaussian loglikelihood function of $\theta = (\beta', \sigma^2, \rho)'$ for the SED model is given by, $\ell_n(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \log|B_n(\rho)| - \frac{1}{2\sigma^2}(Y_n - X_n\beta)'B_n'(\rho)B_n(\rho)(Y_n - X_n\beta)$, where $B_n(\rho) = I_n - \rho W_n$. Maximizing $\ell_n(\theta)$ gives the MLE, $\hat{\theta}_n$ of $\theta$ if the errors are indeed Gaussian, otherwise the QMLE. Given $\rho$, $\ell_n(\theta)$ is partially maximized at,

$$\tilde{\beta}_n(\rho) = [X_n'B_n'(\rho)B_n(\rho)X_n]^{-1}X_n'B_n'(\rho)B_n(\rho)Y_n, \text{ and } \tilde{\sigma}_n^2(\rho) = \tfrac{1}{n}Y_n'B_n'(\rho)M_n(\rho)B_n(\rho)Y_n, \quad (8)$$

where $M_n(\rho) = I_n - B_n(\rho)X_n[X_n'B_n'(\rho)B_n(\rho)X_n]^{-1}X_n'B_n'(\rho)$. The concentrated log-likelihood function for $\rho$ upon substituting the constrained QMLEs $\tilde{\beta}_n(\rho)$ and $\tilde{\sigma}_n^2(\rho)$ into $\ell(\theta)$:

$$\ell_n^c(\rho) = -\frac{n}{2}[\log(2\pi) + 1] + \log|B_n(\rho)| - \frac{n}{2}\log(\tilde{\sigma}_n^2(\rho)). \quad (9)$$

Maximising $\ell_n^c(\rho)$ gives the unconstrained QMLE $\hat{\rho}_n$ of $\rho$, which in turn gives the unconstrained QMLEs of $\beta$ and $\sigma^2$ as, $\hat{\beta}_n \equiv \tilde{\beta}_n(\hat{\rho}_n)$ and $\tilde{\sigma}_n^2 \equiv \hat{\sigma}_n^2(\hat{\rho}_n)$. Thus, $\hat{\theta}_n = (\hat{\beta}_n', \hat{\sigma}_n^2, \hat{\rho}_n)'$.

Liu and Yang (2015) show that, under regularity conditions, the QMLE $\hat{\theta}_n$ is asymptotically normal with mean $\theta_0$, and variance-covariance (VC) matrix $\Sigma_n^{-1}\Gamma_n\Sigma_n^{-1}$, where

$$\Sigma_n = \begin{pmatrix} \frac{1}{\sigma_0^2}X_n'B_n'B_nX_n & 0 & 0 \\ 0 & \frac{n}{2\sigma_0^4} & \frac{1}{\sigma_0^2}\mathrm{tr}(G_n) \\ 0 & \frac{1}{\sigma_0^2}\mathrm{tr}(G_n) & \mathrm{tr}(G_n^sG_n) \end{pmatrix},$$

$$\Gamma_n = \begin{pmatrix} \frac{1}{\sigma_0^2}X_n'B_n'B_nX_n & \frac{1}{2\sigma_0^3}\gamma X_n'B_n'\iota_n & \frac{1}{\sigma_0}\gamma X_n'B_n'g_n \\ \frac{1}{2\sigma_0^3}\gamma\iota_n'B_nX_n & \frac{n}{4\sigma_0^4}(\kappa+2) & \frac{1}{2\sigma_0^2}(\kappa+2)\mathrm{tr}(G_n) \\ \frac{1}{\sigma_0}\gamma g_n'B_nX_n & \frac{1}{2\sigma_0^2}(\kappa+2)\mathrm{tr}(G_n) & \kappa g_n'g_n + \mathrm{tr}(G_n^sG_n) \end{pmatrix},$$

$\iota_n$ is an $n \times 1$ vector of ones, $\gamma$ and $\kappa$ are, respectively, the measures of skewness and excess kurtosis of the idiosyncratic errors $\epsilon_{n,i}$, $g_n = \mathrm{diag}(G_n)$, $G_n = G_n(\rho_0) = W_nB_n^{-1}(\rho_0)$, and $G_n^s = G_n + G_n'$. Based on these results, it is easy to see that $\hat{\beta}_n$ is asymptotically normal with mean $\beta_0$ and variance $\sigma_0^2(X_n'B_n'B_nX_n)^{-1}$. Thus, the inference for $c_0'\beta_0$ is carried out based on the following $t$-ratio:

$$t_{\mathtt{SED}} = \frac{c_0'\hat{\beta}_n - c_0'\beta_0}{\sqrt{\hat{\sigma}_n^2c_0'(X_n'\hat{B}_n'\hat{B}_nX_n)^{-1}c_0}}, \quad (10)$$

where $c_0$ represents a linear contrast of the regression coefficients and $\hat{B}_n = I_n - \hat{\rho}_nW_n$. The $t$-ratio, $t_{\mathtt{SED}}$, is asymptotically $N(0,1)$, and hence inferences concerning $\beta_0$ are carried out by referring to the standard normal critical values.

Liu and Yang (2015) demonstrate based on Monte Carlo experiments that $\hat{\rho}_n$ can be seriously downward biased but the bias of $\hat{\rho}_n$ does not spillover much to $\hat{\beta}_n$. This means that the existence of spatial dependence in the regression errors does not affect much the point estimation of the regression coefficients in terms of consistency and finite sample bias. However, it does spill over to the estimate of $\mathrm{Var}(\hat{\beta}_n)$. First, the downward bias of $\hat{\rho}_n$ causes $\hat{\sigma}_n^2$ to be downward biased when $n$ is not large (e.g., 50). Second, from the expression:

$$X_n'\hat{B}_n'\hat{B}_nX_n = X_n'B_n'B_nX_n - (\hat{\rho}_n-\rho_0)X_n'(W_n'B_n+B_n'W_n)X_n + (\hat{\rho}_n-\rho_0)^2X_n'W_n'W_nX_n,$$

we see that the severe bias of $\hat{\rho}_n$ may cause $X_n'\hat{B}_n'\hat{B}_nX_n$ to be severely biased for the estimation of $X_n'B_n'B_nX_n$. For example when $X_n'(W_n'B_n+B_n'W_n)X_n \geq 0$ (in matrix sense),[1] $X_n'\hat{B}_n'\hat{B}_nX_n$ tends to overestimate $X_n'B_n'B_nX_n$, and hence, $\hat{\sigma}_n^2c_0'(X_n'\hat{B}_n'\hat{B}_nX_n)^{-1}c_0$ tends to underestimate $\mathrm{Var}(c_0'\hat{\beta}_n)$, which makes $t_{\texttt{SED}}$ much more variable than $N(0,1)$ and inferences for $\beta_0$ based on $t_{\texttt{SED}}$ defined in (10) unreliable. Our Monte Carlo results confirm this point.

### 3.2 Bias correction for the SED model

To improve $t_{\texttt{SED}}$, it is necessary to first bias-correct $\hat{\rho}_n$. The method described in Section 2 can be used with $\tilde{\psi}_n(\rho) = \frac{1}{n}\frac{\partial}{\partial\rho}\ell_n^c(\rho)$, where $\ell_n^c(\rho)$ is defined in (9). The following results follow from Liu and Yang (2015):

$$\begin{aligned}
\tilde{\psi}_n(\rho) &= -T_{0n}(\rho) + R_{1n}(\rho), && (11)\\
H_{1n}(\rho) &= -T_{1n}(\rho) + R_{2n}(\rho) + 2R_{1n}^2(\rho), && (12)\\
H_{2n}(\rho) &= -2T_{2n}(\rho) + R_{3n}(\rho) + 6R_{1n}(\rho)R_{2n}(\rho) + 8R_{1n}^3(\rho), && (13)\\
H_{3n}(\rho) &= -6T_{3n}(\rho) + R_{4n}(\rho) + 8R_{1n}(\rho)R_{3n}(\rho) + 6R_{2n}^2(\rho) \\
&\quad +48R_{1n}^2(\rho)R_{2n}(\rho) + 48R_{1n}^4(\rho), && (14)
\end{aligned}$$

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

$$R_{jn}(\rho) = \frac{Y_n'A_n'(\rho)M_n(\rho)D_{jn}(\rho)M_n(\rho)A_n(\rho)Y_n}{Y_n'A_n'(\rho)M_n(\rho)A_n(\rho)Y_n},\ j = 1,2,3,4, \tag{15}$$

with $D_{1n}(\rho) = G_n(\rho)$, and $D_{jn}(\rho), j = 2,3,4,$ being given in Appendix A.

**Bootstrap estimates of biases.** From (11)-(14), we see that $\tilde{\psi}_n$ and $H_{rn}$ are functions of only $R_{jn}, j = 1,\ldots,4$, which are essentially ratios of quadratic forms. Thus, in order to estimate the bias, one needs to estimate the expectations of $R_{jn}$, their powers, cross products, and cross products of powers. It is easy to see that,

$$R_{jn} \equiv R_{jn}(e_n,\rho_0) = \frac{e_n'\Lambda_{jn}(\rho_0)e_n}{e_n'M_n(\rho_0)e_n}, \tag{16}$$

[1] When $W$ follows the `Group Interaction` scheme, this occurs as long as $(\iota_{n_r}'X_{jr})^2 \geq \frac{n_r-1+\rho_0}{(n_r-1)(1-\rho_0)+\rho_0}X_{jr}'X_{jr}$, where $n_r$ is the size of the $r$th group and $X_{jr}$ contains $r$th group values of the $j$th regressor.

where $e_n = \sigma_0^{-1}\epsilon_n$, and $\Lambda_{jn}(\rho_0) = M_n(\rho_0)D_{jn}(\rho_0)M_n(\rho_0)$. It follows that all the necessary quantities whose expectations are required can be expressed in terms of $e_n$ and $\rho_0$, and the general bootstrap procedure described in Section 2 can be followed to give bootstrap estimates of the bias terms $b_{-1}$ and $b_{-3/2}$. See Liu and Yang (2015) for details.

### 3.3 Improved inferences for regression coefficients

First, by simply replacing $\hat\rho$ in $t_{\texttt{SED}}$ defined in (10) by $\hat\rho_n^{\rm bc2}$, the second-order bias-corrected $\hat\rho$, we obtain the following potentially improved statistic:

$$t_{\texttt{SED}}^{\rm bc} = \frac{c_0'\hat\beta_n^{\rm bc} - c_0'\beta_0}{\sqrt{\hat\sigma_n^{2,\rm bc}c_0'(X_n'\hat B_n^{\rm bc2\prime}\hat B_n^{\rm bc2}X_n)^{-1}c_0}}, \tag{17}$$

where $\hat\beta_n^{\rm bc} = \tilde\beta_n(\hat\rho_n^{\rm bc2})$, $\hat\sigma_n^{2,\rm bc} = \tilde\sigma_n^2(\hat\rho_n^{\rm bc2})$, and $\hat B_n^{\rm bc2} = I_n - \hat\rho_n^{\rm bc2}W_n$. Obviously, this statistic is not fully second-order bias-corrected. However, Monte Carlo results presented in the next subsection show that it offers a huge improvement over $t_{\texttt{SED}}$. This confirms the point made at the end of Section 3.1. However, results also show that when $n$ is not so large, there is still room for further improvement on $t_{\texttt{SED}}^{\rm bc}$.

Let $F_n(\rho) = [X_n'B_n'(\rho)B_n(\rho)X_n]^{-1}X_n'B_n'(\rho)B_n(\rho)$ such that $\tilde\beta_n(\rho) = F_n(\rho)Y_n$ defined in (8) and denoting $\tilde\beta_n = \tilde\beta_n(\rho_0)$, and $\tilde\beta_n^{(r)} = \frac{d^r}{d\rho_0^r}\tilde\beta_n(\rho_0)$ and $F_n^{(r)} = \frac{d^r}{d\rho_0^r}F_n(\rho_0)$ for $r = 1, 2$, we have the following second-order stochastic expansion for $\hat\beta_n = \tilde\beta_n(\hat\rho_n)$:

$$\begin{aligned}\hat\beta_n - \beta_0 &= \tilde\beta_n - \beta_0 + \tilde\beta_n^{(1)}(\hat\rho_n - \rho_0) + \tfrac{1}{2}\tilde\beta_n^{(2)}(\hat\rho_n - \rho_0)^2 + O_p(n^{-3/2})\\ &= b_{0n} + \mathrm{E}(\tilde\beta_n^{(1)})(a_{-1/2} + a_{-1}) + b_{1n}a_{-1/2} + \tfrac{1}{2}\mathrm{E}(\tilde\beta_n^{(2)})a_{-1/2}^2 + O_p(n^{-3/2}),\end{aligned} \tag{18}$$

where $b_{0n} = F_nB_n^{-1}\epsilon_n$, $b_{1n} = F_n^{(1)}B_n^{-1}\epsilon_n$, $\mathrm{E}(\tilde\beta_n^{(1)}) = F_n^{(1)}X_n\beta_0$, $\mathrm{E}(\tilde\beta_n^{(2)}) = F_n^{(2)}X_n\beta_0$, and $F_n^{(r)}$ are given in Appendix A. This leads immediately to, as a by-product of the bootstrap bias-correction for $\hat\rho_n$, a fully 2nd-order bias-corrected estimator $\hat\beta_n^{\rm bc2}$ of $\beta$. Similarly, an expansion as (18) can easily be carried out for $\hat\sigma_n^2 = \tilde\sigma_n^2(\hat\rho_n)$, giving a fully 2nd-order bias-corrected estimator $\hat\sigma_n^{2,\rm bc2}$ of $\sigma^2$.[2] Finally, denoting $g(e_n, \theta_0) \equiv b_{0n} + \mathrm{E}(\tilde\beta_n^{(1)})(a_{-1/2} + a_{-1}) + b_{1n}a_{-1/2} + \frac{1}{2}\mathrm{E}(\tilde\beta_n^{(2)})a_{-1/2}^2$, the expansion (18) leads to a second-order variance expansion:

$$\mathrm{Var}(\hat\beta_n) = \mathrm{Var}[g(e_n, \theta_0)] + O(n^{-2}).$$

Further it is easy to see $\mathrm{Var}(\hat\beta_n^{\rm bc2}) = \mathrm{Var}(\hat\beta_n) + O(n^{-2})$, and $\mathrm{Var}(\hat\beta_n^{\rm bc}) = \mathrm{Var}(\hat\beta_n) + O(n^{-2})$. Obviously, an explicit expression of the above is difficult to obtain, but is not needed as it can be easily estimated by the two-stage bootstrap procedure described below. Recall $a_{-1/2} = \Omega_n\tilde\psi_n$, and $a_{-1} = \Omega_nH_{1n}^\circ a_{-1/2} + \frac{1}{2}\Omega_n\mathrm{E}(H_{2n})(a_{-1/2}^2) = \Omega_n\tilde\psi_n + \Omega_n^2H_{1n}\tilde\psi_n + \frac{1}{2}\Omega_n^3\mathrm{E}(H_{2n})\tilde\psi_n^2$.

[2] As $\hat\beta_n^{\rm bc}$ and $\hat\beta_n^{\rm bc2}$ do not differ much, and $\hat\sigma_n^{2,\rm bc}$ and $\hat\sigma_n^{2,\rm bc2}$ also do not differ much, one can simply use $\hat\beta_n^{\rm bc}$ and $\hat\sigma_n^{2,\rm bc}$ in practical applications.

`Stage 1:` Compute $\hat{\theta}_n$ and the QML residuals $\hat{e}_n = \hat{\sigma}_n^{-1}\hat{B}_n(Y_n - X_n\hat{\beta}_n)$. Resample $\hat{e}_n$ to give $\hat{\rho}_n^{\mathrm{bc2}}$, and hence $\hat{\beta}_n^{\mathrm{bc2}}$ and $\hat{\sigma}_n^{2,\mathrm{bc2}}$, using the algorithm `BA-1` given in Section 2.

`Stage 2:` Update the QML residuals as $\hat{e}_n^{\mathrm{bc2}} = \hat{\sigma}_n^{\mathrm{bc2},-1}\hat{B}_n^{\mathrm{bc2}}(Y_n - X_n\hat{\beta}_n^{\mathrm{bc2}})$ and compute $g_{n,b}^* \equiv g(\hat{e}_{n,b}^{\mathrm{bc2}*}, \hat{\theta}_n^{\mathrm{bc2}})$ for $b = 1, \ldots, B$, where $\hat{e}_{n,b}^{\mathrm{bc2}*}$ is the $b$th bootstrap sample drawn from the EDF of $\hat{e}_n^{\mathrm{bc2}}$, and $\hat{\theta}_n^{\mathrm{bc2}} = (\hat{\beta}_n^{\mathrm{bc2}\prime}, \hat{\sigma}_n^{\mathrm{bc2}}, \hat{\rho}_n^{\mathrm{bc2}})'$. The bootstrap estimate of $\mathrm{Var}(\hat{\beta}_n^{\mathrm{bc2}})$, unbiased up to $O(n^{-3/2})$, is thus, $\widehat{\mathrm{Var}}(\hat{\beta}_n^{\mathrm{bc2}}) = \frac{1}{B}\sum_{b=1}^B g_{n,b}^* g_{n,b}^{*\prime} - \frac{1}{B}\sum_{b=1}^B g_{n,b}^* \frac{1}{B}\sum_{b=1}^B g_{n,b}^{*\prime}$.

We have a second-order 'bias-corrected' $t$-statistic as follows:

$$t_{\mathtt{SED}}^{\mathrm{bc2}} = \frac{c_0'\hat{\beta}_n^{\mathrm{bc2}} - c_0'\beta_0}{\sqrt{c_0'\widehat{\mathrm{Var}}(\hat{\beta}_n^{\mathrm{bc2}})c_0}}. \tag{19}$$

## 3.4 Monte Carlo results

Finite sample performance of $t_{\mathtt{SED}}$, $t_{\mathtt{SED}}^{\mathrm{bc}}$ and $t_{\mathtt{SED}}^{\mathrm{bc2}}$ is investigated and compared under the following data generating process (DGP):

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

where $X_{1n}$ and $X_{2n}$ are the $n \times 1$ vectors containing the values of two fixed regressors. The parameters of the simulation are initially set to be as: $\beta = (5, 1, 1)'$, $\sigma^2 = 1$, $\rho$ takes values form $\{-0.5, -0.25, 0, 0.25, 0.5\}$ and $n$ take values from $\{50, 100, 200, 500\}$. Each set of Monte Carlo results is based on $M = 10,000$ Monte Carlo samples, and $B = 999 + \lfloor n^{0.75} \rfloor$ bootstrap samples within each Monte Carlo sample. The methods for generating $X_n$, $W_n$, and the errors are described in Appendix A.

Table 3.1 summarizes some results for $t_{\mathtt{SED}}$ and $t_{\mathtt{SED}}^{\mathrm{bc2}}$ used for testing $H_0 : \beta_1 = \beta_2$. From the results we see that (i) as $n$ increases, all tests converge in terms of rejection rates, (ii) it is indeed the case that the asymptotic test $t_n$ can be very unreliable in the sense it rejects the true $H_0$ much too often than it supposes to. The test $t_{\mathtt{SED}}^{\mathrm{bc}}$ offers a huge reduction in size distortions, and when $n = 200$ and 500, its rejection rates become very close to their nominal levels. Nevertheless, when $n = 50$ or 100, we see from the tables that there is room for further improvement on $t_{\mathtt{SED}}^{\mathrm{bc}}$. The $t$-statistic $t_{\mathtt{SED}}^{\mathrm{bc2}}$ based on the second order corrected variance provides a further improvement on $t_{\mathtt{SED}}^{\mathrm{bc}}$ with the rejection rates quite close to the nominal levels even when $n$ is not so large. The results show that the error distribution does not significantly affect the performance of the three tests. The true value of the spatial parameter has little effect on the performance of the two improved tests (except when $n = 50$), but has a significant effect on the asymptotic test: the size distortion gets larger when $\rho$ changes from .5 to $-.5$. Furthermore, the size distortion for the asymptotic test is seen to be quite persistent, which remains to be at least 20% even when $n = 500$. The results (unreported for brevity) show that the tests under a more sparse spatial weight matrix generally have smaller size distortions.

# 4. Improved Inferences for the SLD Model

This section concerns the improved inference methods for the regression coefficients of the SLD model. Section 4.1 outlines the asymptotic results, and Section 4.2 the finite sample bias-correction results. Section 4.3 presents the improved inference methods, and Section 4.4 presents Monte Carlo results.

## 4.1 QML estimation and asymptotic inference

The regression model with spatial lag dependence (SLD) takes the form:

$$Y_n = \lambda W_n Y_n + X_n\beta + \epsilon_n, \tag{20}$$

Letting $A_n(\lambda) = I_n - \lambda W_n$, the log-likelihood function of $\theta = (\beta', \sigma^2, \lambda)'$ is $\ell_n(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \log|A_n(\lambda)| - \frac{1}{2\sigma^2}\left[A_n(\lambda)Y_n - X_n\beta\right]'\left[A_n(\lambda)Y_n - X_n\beta\right]$. Given $\lambda$, $\ell_n(\theta)$ is maximized at

$$\tilde{\beta}_n(\lambda) = (X_n'X_n)^{-1}X_n'A_n(\lambda)Y_n \quad \text{and} \quad \tilde{\sigma}_n^2(\lambda) = \tfrac{1}{n}Y_n'A_n'(\lambda)M_nA_n(\lambda)Y_n, \tag{21}$$

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

$$\ell_n^c(\lambda) = -\frac{n}{2}[\log(2\pi)+1] - \frac{n}{2}\log\hat{\sigma}_n^2(\lambda) + \log|A_n(\lambda)|. \tag{22}$$

Maximizing $\ell_n^c(\lambda)$ gives the unconstrained QMLE $\hat{\lambda}_n$ of $\lambda$. The unconstrained QMLEs of $\beta$ and $\sigma^2$ are thus, $\hat{\beta}_n \equiv \tilde{\beta}_n(\hat{\lambda}_n)$ and $\hat{\sigma}_n^2 \equiv \tilde{\sigma}_n^2(\hat{\lambda}_n)$. Write $\hat{\theta}_n = (\hat{\beta}_n', \hat{\sigma}_n^2, \hat{\lambda}_n)'$.

Lee (2004) shows that $\hat{\theta}_n$ is asymptotically $N(\theta_0, \Sigma_n^{-1}\Gamma_n\Sigma_n^{-1})$, where

$$\Sigma_n = \begin{pmatrix} \frac{1}{\sigma_0^2}X_n'X_n & 0 & \frac{1}{\sigma_0}X_n'\eta_n \\ 0 & \frac{n}{2\sigma_0^4} & \frac{1}{\sigma_0^2}\mathrm{tr}(G_n) \\ \frac{1}{\sigma_0}\eta_n'X_n & \frac{1}{\sigma_0^2}\mathrm{tr}(G_n) & \eta_n'\eta_n + \mathrm{tr}(G_n^sG_n) \end{pmatrix},$$

$$\Gamma_n = \begin{pmatrix} 0 & \frac{1}{2\sigma_0^3}\gamma X_n'\iota_n & \frac{1}{\sigma_0}\gamma X_n'g_n \\ \frac{1}{2\sigma_0^3}\gamma\iota_n'X_n & \frac{n}{4\sigma_0^4}\kappa & \frac{1}{2\sigma_0^2}\gamma\iota_n'\eta_n + \frac{1}{2\sigma_0^2}\kappa\mathrm{tr}(G_n) \\ \frac{1}{\sigma_0}\gamma g_n'X_n & \frac{1}{2\sigma_0^2}\gamma\iota_n'\eta_n + \frac{1}{2\sigma_0^2}\kappa\mathrm{tr}(G_n) & \kappa g_n'g_n + 2\gamma g_n'\eta_n \end{pmatrix} + \Sigma_n,$$

$\iota_n$, $\gamma$ and $\kappa$ are defined as in Section 3.1, $g_n = \mathrm{diag}(G_n)$, $G_n = G_n(\rho_0) = W_nA_n^{-1}(\lambda_0)$, $G_n^s = G_n + G_n'$ and $\eta_n = \sigma_0^{-1}G_nX_n\beta_0$.

Letting $V_{n1}$ be the submatrix of $\Sigma_n^{-1}\Gamma_n\Sigma_n^{-1}$ corresponding to $\beta$ and $\hat{V}_{n1}$ be its estimate, an asymptotic $t$-statistic for inferences for $c_0'\beta_0$ is thus,

$$t_{\mathtt{SLD}} = \frac{c_0'\hat{\beta}_n - c_0'\beta_0}{\sqrt{c_0'\hat{V}_{n1}c_0}}, \tag{23}$$

which is asymptotically $N(0,1)$. Finite sample properties of $t_{\texttt{SLD}}$ is of interest.

As $\tilde{\beta}_n(\hat{\lambda}) = \beta_0 + (\lambda - \hat{\lambda})(X_n'X_n)^{-1}X_n'G_nX_n\beta_0 + o_p(1)$, we see clearly that any estimation bias of $\hat{\lambda}$ is quickly passed down to the QMLE of $\beta_0$. Thus the $t$-statistic computed using $\tilde{\beta}_n(\hat{\lambda})$ and the variance estimate $\hat{V}_{n1}$ can be unreliable. This fact is confirmed by the Monte Carlo results. As such it is desirable to find ways to improve $t_{\texttt{SLD}}$.

## 4.2 Bias corrections

As an illustration to his general bias correction method, Yang (2015b) studied the SLD model in detail. Letting $\tilde{\psi}_n(\lambda) = \frac{\partial}{\partial\lambda}\ell_n^c(\lambda)$, where $\ell_n^c(\lambda)$ is given in (22), we have,

$$\begin{aligned}
\tilde{\psi}_n(\lambda) &= -h_nT_{0n}(\lambda) + h_nR_{1n}(\lambda), &(24)\\
H_{1n}(\lambda) &= -T_{1n}(\lambda) - R_{2n}(\lambda) + 2R_{1n}^2(\lambda), &(25)\\
H_{2n}(\lambda) &= -2T_{2n}(\lambda) - 6R_{1n}(\lambda)R_{2n}(\lambda) + 8R_{1n}^3(\lambda), &(26)\\
H_{3n}(\lambda) &= -6T_{3n}(\lambda) + 6R_{2n}^2(\lambda) - 48R_{1n}^2(\lambda)R_{2n}(\lambda) + 48R_{1n}^4(\lambda), &(27)
\end{aligned}$$

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

$$R_{1n}(\lambda) = \frac{Y_n'A_n'(\lambda)M_nW_nY_n}{Y_n'A_n'(\lambda)M_nA_n(\lambda)Y_n} \quad \text{and} \quad R_{2n}(\lambda) = \frac{Y_n'W_n'M_nW_nY_n}{Y_n'A_n'(\lambda)M_nA_n(\lambda)Y_n}. \tag{28}$$

**Bootstrap estimates of biases.** The two key ratios can be written as:

$$\begin{aligned}
R_{1n}(e_n,\theta_0) &= \frac{e_n'M_nG_ne_n + e_n'M_n\eta_n}{e_n'M_ne_n},\\
R_{2n}(e_n,\theta_0) &= \frac{e_n'G_n'M_nG_ne_n + 2e_n'G_n'M_n\eta_n + \eta_n'M_n\eta_n}{e_n'M_ne_n},
\end{aligned}$$

where $e_n = \sigma_0^{-1}\epsilon_n$. Hence, $\tilde{\psi}_n = \tilde{\psi}_n(e_n,\theta_0)$ and $H_{rn} = H_{rn}(e_n,\theta_0)$ $r = 1,2,3$. In other words, all the random quantities in the bias term can be expressed in terms of $e_n$ and $\theta_0$. So, the bias corrections are carried out in a similar manner using an estimate of $R_{1n}(e_n,\theta_0)$ and $R_{2n}(e_n,\theta_0)$. See Yang (2015b) for details. Let $\hat{\lambda}_n^{\text{bc2}}$ be the second-order bias corrected $\hat{\lambda}_n$, and let $\hat{\beta}_n^{\text{bc}} = \tilde{\beta}(\hat{\lambda}_n^{\text{bc2}})$ and $\hat{\sigma}_n^{2,\text{bc}} = \tilde{\sigma}_n^2(\hat{\lambda}_n^{\text{bc2}})$.

## 4.3 Improved inferences for regression coefficients

Similar to the case of SED model, replacing $\hat{\lambda}_n$ by $\hat{\lambda}_n^{\text{bc2}}$ in the definition of $t_{\texttt{SLD}}$, we obtain a statistic which is expected to have a better finite sample performance:

$$t_{\texttt{SLD}}^{\text{bc}} = \frac{c_0'\hat{\beta}_n^{\text{bc}} - c_0'\beta_0}{\sqrt{c_0'\hat{V}_{n1}^{\text{bc}}c_0}}, \tag{29}$$

where $\hat{V}_{n1}^{\text{bc}}$ is $V_{n1}$ evaluated at $\hat{\lambda}_n^{\text{bc2}}$, $\hat{\beta}_n^{\text{bc}}$, $\hat{\sigma}_n^{2,\text{bc}}$, $\hat{\gamma}_n^{\text{bc}}$, and $\hat{\kappa}_n^{\text{bc}}$. The last two are the estimates of $\gamma$ and $\kappa$, the skewness and excess kurtosis of $\epsilon_{n,i}$ involved in $V_{n1}$.

Now, to further improve $t_{\texttt{SLD}}^{\text{bc}}$, note that

$$\begin{aligned}\hat{\beta}_n - \beta_0 &= \tilde{\beta}_n - \beta_0 - (\hat{\lambda}_n - \lambda_0)(X_n'X_n)^{-1}X_n'G_nX_n\beta_0 - (\hat{\lambda}_n - \lambda_0)(X_n'X_n)^{-1}X_n'G_n\epsilon_n \\ &= (X_n'X_n)^{-1}X_n'\big[\epsilon_n - (a_{-1/2} + a_{-1})G_nX_n\beta_0 - a_{-1/2}G_n\epsilon_n\big] + O_p(n^{-3/2}).\end{aligned} \tag{30}$$

This leads immediately to a 2nd-order bias-corrected estimator $\hat{\beta}_n^{\text{bc2}}$ of $\beta$, and a second-order expansion for $\text{Var}(\hat{\beta}_n)$ as,

$$\text{Var}(\hat{\beta}_n) = (X_n'X_n)^{-1}X_n'\text{Var}\big[\epsilon_n - (a_{-1/2} + a_{-1})G_nX_n\beta_0 - a_{-1/2}G_n\epsilon_n\big]X_n(X_n'X_n)^{-1} + O(n^{-2}).$$

It is easy to see $\text{Var}(\hat{\beta}_n^{\text{bc2}}) = \text{Var}(\hat{\beta}_n) + O(n^{-2})$. As in (30), an expansion can be carried out for $\hat{\sigma}_n^2$ in terms of $\hat{\lambda}_n$, leading to a 2nd-order bias-corrected estimator $\sigma_n^{2,\text{bc2}}$ of $\sigma^2$.[3] Similarly, a two-stage bootstrap procedure can be followed to give a consistent estimate of $V = X_n'\text{Var}\big[\epsilon_n - (a_{-1/2} + a_{-1})G_nX_n\beta_0 - a_{-1/2}G_n\epsilon_n\big]X_n$: first, run the algorithm `BA-1` to give 2nd-order bias-corrected estimators $\hat{\lambda}_n^{\text{bc2}}$, $\hat{\beta}_n^{\text{bc2}}$ and $\sigma_n^{2,\text{bc2}}$; then update the residuals and run the algorithm `BA-1` again using the updated residuals to give a sequence of bootstrap values for $V$, and hence the bootstrap estimate $\widehat{\text{Var}}(\hat{\beta}_n^{\text{bc2}})$ of $\text{Var}(\hat{\beta}_n)$. The resulted 2nd-order bias-corrected $t$-statistic is thus:

$$t_{\texttt{SLD}}^{\text{bc2}} = \frac{c_0'\hat{\beta}_n^{\text{bc2}} - c_0'\beta_0}{\sqrt{c_0'\widehat{\text{Var}}(\hat{\beta}_n^{\text{bc2}})c_0}}. \tag{31}$$

### 4.4 Monte Carlo results

Finite sample performance of $t_{\texttt{SLD}}$, $t_{\texttt{SLD}}^{\text{bc}}$ and $t_{\texttt{SLD}}^{\text{bc2}}$ is investigated under the following DGP:

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

where all the quantities are generated in a similar manner to those for the SED model. The parameters for the Monte Carlo simulation are also set to be the same values as before.

Table 4.1 summarizes some empirical sizes of the tests $t_{\texttt{SLD}}$, $t_{\texttt{SLD}}^{\text{bc}}$ and $t_{\texttt{SLD}}^{\text{bc2}}$ when used for testing $H_0 : \beta_1 = \beta_2$ under the `Group Interaction` scheme. From the results we see that (i) as $n$ increases, all tests converge in terms of sizes, (ii) it is indeed the case that the asymptotic test $t_{\texttt{SLD}}$ can be very unreliable in the sense that it rejects the true $H_0$ much too often than it supposes to. The test $t_{\texttt{SLD}}^{\text{bc2}}$ offers a huge reduction in size distortions, with the empirical sizes getting close to their nominal levels faster than in the SED case. Nevertheless, when $n = 50$, the results show that $t_{\texttt{SLD}}^{\text{bc}}$ needs further improvements, and indeed the test $t_{\texttt{SLD}}^{\text{bc2}}$ based on the

[3]Again, the estimators $\hat{\beta}_n^{\text{bc}}$ and $\hat{\beta}_n^{\text{bc2}}$ do not differ much, and the estimators $\hat{\sigma}_n^{2,\text{bc}}$ and $\hat{\sigma}_n^{2,\text{bc2}}$ do not differ much. Hence in practical applications, one can use the simpler versions $\hat{\beta}_n^{\text{bc}}$ and $\hat{\sigma}_n^{2,\text{bc}}$.

second-order corrected variance offers the desired improvements. More results are available from the authors upon request.

# 5. Improved Inferences for the SARAR Model

In this section, we study the finite sample bias problem of the general SARAR model and introduce improved inference methods for the regression coefficients of this model. Neither issue has been formally considered due to its complexity, and hence the results presented in this section constitute important contributions to the literature, in particular considering the fact that the SARAR model is more versatile and hence practically more useful than either the SLD model or the SED model. Section 5.1 outlines the QML estimation and the asymptotic inference method. Section 5.2 presents detailed results for bias-correcting the QMLEs of the spatial parameters. Section 5.3 presents the improved inference methods for the regression coefficients. Section 5.4 presents Monte Carlo results.

## 5.1 QML estimation and asymptotic inference

Combining the SED and SLD models considered above, we have the so-called spatial autoregressive model with autoregressive errors, also known as the SARAR model:

$$Y_n = \lambda W_{1n} Y_n + X_n \beta + u_n, \quad u_n = \rho W_{2n} u_n + \epsilon_n. \tag{32}$$

The Gaussian log-likelihood function of $\theta = (\beta', \sigma^2, \lambda, \rho)'$ is $\ell_n(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \log|A_n(\lambda)| + \log|B_n(\rho)| - \frac{1}{2\sigma^2}[Y_n(\delta) - X_n(\rho)\beta]'[Y_n(\delta) - X_n(\rho)\beta]$, where $A_n(\lambda) = I_n - \lambda W_{1n}$, $B_n(\rho) = I_n - \rho W_{2n}$, $X_n(\rho) = B_n(\rho)X_n$ and $Y_n(\delta) = B_n(\rho)A_n(\lambda)Y_n$. The constrained QMLEs of $\beta$ and $\sigma^2$, given $\delta = (\lambda, \rho)'$, are

$$\tilde{\beta}_n(\delta) = [X_n'(\rho)X_n(\rho)]^{-1}X_n'(\rho)Y_n(\delta) \ \text{ and } \ \tilde{\sigma}_n^2(\delta) = \tfrac{1}{n}Y_n'(\delta)M_n(\rho)Y_n(\delta), \tag{33}$$

where $M_n(\rho) = I_n - B_n(\rho)X_n[X_n'B_n'(\rho)B_n(\rho)X_n]^{-1}X_n'B_n'(\rho)$. Then, the concentrated Gaussian loglikelihood function for $\delta$ is,

$$\ell_n^c(\delta) = -\frac{n}{2}[\ln(2\pi) + 1] - \frac{n}{2}\ln(\hat{\sigma}_n^2(\delta)) + \ln|A_n(\lambda)| + \ln|B_n(\rho)|. \tag{34}$$

Maximizing (34) gives the QMLE $\hat{\delta}_n$ of $\delta$, and thus the QMLEs of $\beta$ and $\sigma^2$ as $\hat{\beta}_n \equiv \tilde{\beta}_n(\hat{\delta}_n)$ and $\hat{\sigma}_n^2 \equiv \tilde{\sigma}_n^2(\hat{\delta}_n)$. Write $\hat{\theta}_n = (\hat{\beta}_n', \hat{\sigma}_n^2, \hat{\delta}_n')'$. The concentrated score function upon dividing by $n$ is,

$$\tilde{\psi}_n(\delta) = \begin{cases} -\dfrac{1}{n}\mathrm{tr}(G_{1n}(\lambda)) + \dfrac{Y_n'(\delta)M_n(\rho)\bar{B}_n(\delta)Y_n(\delta)}{Y_n'(\delta)M_n(\rho)Y_n(\delta)}, \\ -\dfrac{1}{n}\mathrm{tr}(G_{2n}(\rho)) + \dfrac{Y_n'(\delta)M_n(\rho)G_{2n}(\rho)M_n(\rho)Y_n(\delta)}{Y_n'(\lambda)M_n(\rho)Y_n(\delta)}, \end{cases} \tag{35}$$

where $G_{1n}(\lambda) = W_{1n}A_n^{-1}(\lambda)$, $G_{2n}(\rho) = W_{2n}B_n^{-1}(\rho)$ and $\bar{B}_n(\delta) = B_n(\rho)G_{1n}(\lambda)B_n^{-1}(\rho)$.

Jin and Lee (2013) shows that under some regularity conditions, $\hat{\theta}_n$ is asymptotically normal with mean $\theta_0$ and asymptotic VC matrix $\Sigma_n^{-1}\Gamma_n\Sigma_n^{-1}$, where

$$\Sigma_n = \begin{pmatrix} \frac{1}{\sigma_0^2}X_n'B_n'B_nX_n & 0 & \frac{1}{\sigma_0}X_n'B_n'\mu_n & 0 \\ 0 & \frac{n}{2\sigma_0^4} & \frac{1}{\sigma_0^2}\mathrm{tr}(\bar{B}_n) & \frac{1}{\sigma_0^2}\mathrm{tr}(G_{2n}) \\ \frac{1}{\sigma_0}\mu_n'B_nX_n & \frac{1}{\sigma_0^2}\mathrm{tr}(\bar{B}_n) & \mu_n'\mu_n + \mathrm{tr}(\bar{B}_n^s\bar{B}_n) & \mathrm{tr}(G_{2n}^s\bar{B}_n) \\ 0 & \frac{1}{\sigma_0^2}\mathrm{tr}(G_{2n}) & \mathrm{tr}(G_{2n}^s\bar{B}_n) & \mathrm{tr}(G_{2n}^sG_{2n}) \end{pmatrix},$$

$$\Gamma_n = \begin{pmatrix} 0 & \frac{\gamma}{2\sigma_0^3}X_n'B_n'\iota_n & \frac{\gamma}{\sigma_0}X_n'B_n'\bar{b}_n & \frac{\gamma}{\sigma_0}X_n'B_n'g_{2n} \\ \frac{\gamma}{2\sigma_0^3}\iota_n'B_nX_n & \frac{n\kappa}{4\sigma_0^4} & \frac{\kappa}{2\sigma_0^2}\mathrm{tr}(\bar{B}_n) + \frac{\gamma}{2\sigma_0^2}\iota_n'\mu_n & \frac{\kappa}{2\sigma_0^2}\mathrm{tr}(G_{2n}) \\ \frac{\gamma}{\sigma_0}\bar{b}_n'B_nX_n & \frac{\kappa}{2\sigma_0^2}\mathrm{tr}(\bar{B}_n) + \frac{\gamma}{2\sigma_0^2}\iota_n'\mu_n & \kappa\bar{b}_n'\bar{b}_n + 2\gamma\bar{b}_n'\mu_n & \kappa g_{2n}'\bar{b}_n + \gamma g_{2n}'\mu_n \\ \frac{\gamma}{\sigma_0}g_{2n}'B_nX_n & \frac{\kappa}{2\sigma_0^2}\mathrm{tr}(G_{2n}) & \kappa g_{2n}'\bar{b}_n + \gamma g_{2n}'\mu_n & \kappa g_{2n}'g_{2n} \end{pmatrix} + \Sigma_n,$$

$\iota_n$, $\gamma$ and $\kappa$ are defined in the earlier sections, $\mu_n = \sigma_0^{-1}B_nG_{1n}X_n\beta_0$, $\bar{b}_n = \mathrm{diag}(\bar{B}_n)$, $g_{2n} = \mathrm{diag}(G_{2n})$, $\bar{B}_n^s = \bar{B}_n + \bar{B}_n'$ and $G_{2n}^s = G_{2n} + G_{2n}'$.

Letting $V_{n1}$ be the submatrix of $\Sigma_n^{-1}\Gamma_n\Sigma_n^{-1}$ corresponding to $\beta$ and $\hat{V}_{n1}$ be its estimate, an asymptotic $t$-statistic for inferences for $c_0'\beta_0$ is thus,

$$t_{\texttt{SARAR}} = \frac{c_0'\hat{\beta}_n - c_0'\beta_0}{\sqrt{c_0'\hat{V}_{n1}c_0}} \sim N(0,1). \tag{36}$$

As indicated in the introduction, there is no formal treatment in the literature in terms of the finite sample bias of the QML estimators of the SARAR model. Given the fact that the QMLEs of the spatial parameters in the SED and SLD models can both be seriously biased, there is a good reason to believe that they will remain to be biased when the spatial effects are combined. Hence bias corrections for the QMLEs of the SARAR model would again be useful in improving the inference methods for the model.

## 5.2 Bias corrections

Bias correction can be carried out as an application to the general method of bias correction of Yang (2015b), for a vector of nonlinear estimators. To do so we need the higher-order partial derivatives of $\tilde{\psi}_n(\delta)$, $H_{rn}(\delta) = \nabla^r\tilde{\psi}_n(\delta), r = 1,2,3$, where the partial derivatives are obtained sequentially and elementwise with respect to $\delta'$. Define, $T_{rn} = \mathrm{tr}(G_{1n}^r(\lambda))$ and $K_{rn} = \mathrm{tr}(G_{2n}^r(\rho))$, $r = 0,1,2,3$. Also define the following quantities,

$$R_{1n}(\delta) = \frac{Y_n'(\delta)M_n(\rho)\bar{B}_n(\delta)Y_n(\delta)}{Y_n'(\delta)M_n(\rho)Y_n(\delta)},$$

$$R_{2n}(\delta) = \frac{Y_n'(\delta)\bar{B}_n'(\delta)M_n(\rho)\bar{B}_n(\delta)Y_n(\delta)}{Y_n'(\delta)M_n(\rho)Y_n(\delta)},$$

$$S_{rn}(\delta) = \frac{Y_n'(\delta)M_n(\rho)D_{rn}(\rho)M_n(\rho)Y_n(\delta)}{Y_n'(\delta)M_n(\rho)Y_n(\delta)},\ r=1,2,3,4,$$

$$Q^{\dagger}_{rn}(\delta) = \frac{Y_n'(\delta)M_n(\rho)D_{rn}(\rho)M_n(\rho)\bar{B}_n(\delta)Y_n(\delta)}{Y_n'(\delta)M_n(\rho)Y_n(\delta)},\ r=1,2,3,$$

$$Q^{\ddagger}_{rn}(\delta) = \frac{Y_n'(\delta)\bar{B}_n'(\delta)M_n(\rho)D_{rn}(\rho)M_n(\rho)\bar{B}_n(\delta)Y_n(\delta)}{Y_n'(\delta)M_n(\rho)Y_n(\delta)},\ r=1,2,$$

where $D_{1n}(\rho) = G_{2n}(\rho)$, and $D_{rn}(\rho)$, $r = 2,3,4$, are given in Appendix A. These quantities have the following properties,

$$\begin{aligned}
&\tfrac{d}{d\lambda}R_{1n}(\delta) = 2R^2_{1n}(\delta) - R_{2n}(\delta), && \tfrac{d}{d\lambda}R_{2n}(\delta) = 2R_{1n}(\delta)R_{2n}(\delta),\\
&\tfrac{d}{d\lambda}S_{rn}(\delta) = 2R_{1n}(\delta)S_{rn}(\delta) - 2Q^{\dagger}_{rn}(\delta), && \tfrac{d}{d\lambda}Q^{\dagger}_{rn}(\delta) = 2R_{1n}(\delta)Q^{\dagger}_{rn}(\delta) - Q^{\ddagger}_{rn}(\delta),\\
&\tfrac{d}{d\lambda}Q^{\ddagger}_{rn}(\delta) = 2R_{1n}(\delta)Q^{\ddagger}_{rn}(\delta), &&\\
&\tfrac{d}{d\rho}R_{1n}(\delta) = 2R_{1n}(\delta)S_{1n}(\delta) - 2Q^{\dagger}_{1n}(\delta), && \tfrac{d}{d\rho}R_{2n}(\delta) = 2R_{2n}(\delta)S_{1n}(\delta) - 2Q^{\ddagger}_{1n}(\delta),\\
&\tfrac{d}{d\rho}S_{rn}(\delta) = 2S_{1n}(\delta)S_{rn}(\delta) + S_{r+1,n}(\delta), && \tfrac{d}{d\rho}Q^{\dagger}_{rn}(\delta) = 2S_{1n}(\delta)Q^{\dagger}_{rn}(\delta) + Q^{\dagger}_{r+1,n}(\delta),\\
&\tfrac{d}{d\rho}Q^{\ddagger}_{rn}(\delta) = 2S_{1n}(\delta)Q^{\ddagger}_{rn}(\delta) + Q^{\ddagger}_{r+1,n}(\delta). &&
\end{aligned}$$

Write $\tilde{\psi}_n(\delta) = (\tilde{\psi}_{1n}(\delta),\ \tilde{\psi}_{2n}(\delta))'$, where $\tilde{\psi}_{1n}(\delta) = -T_{0n}(\lambda) + R_{1n}(\delta)$ and $\tilde{\psi}_{2n}(\delta) = -K_{0n}(\rho) + S_{1n}(\delta)$. Denote the partial derivatives of $\tilde{\psi}_{rn}(\delta)$ by adding superscripts $\lambda$ and/or $\rho$ sequentially, e.g., $\tilde{\psi}^{\lambda\lambda}_{1n}(\delta) = \frac{\partial^2}{\partial\lambda^2}\tilde{\psi}_{1n}(\delta)$, and $\tilde{\psi}^{\lambda\rho\lambda}_{2n}(\delta) = \frac{\partial^3}{\partial\lambda\partial\rho\partial\lambda}\tilde{\psi}_{2n}(\delta)$. Thus, $H_{1n}(\delta)$ has 1st row $\{\tilde{\psi}^{\lambda}_{1n}(\delta), \tilde{\psi}^{\rho}_{1n}(\delta)\}$ and 2nd row $\{\tilde{\psi}^{\lambda}_{2n}(\delta), \tilde{\psi}^{\rho}_{2n}(\delta)\}$, which gives,

$$H_{1n}(\delta) = \begin{pmatrix} -T_{1n}(\lambda) - R_{2n}(\delta) + 2R^2_{1n}(\delta), & -2Q^{\dagger}_{1n}(\delta) + 2R_{1n}(\delta)S_{1n}(\delta) \\ -2Q^{\dagger}_{1n}(\delta) + 2R_{1n}(\delta)S_{1n}(\delta), & -K_{1n}(\rho) + S_{2n}(\delta) + 2S^2_{1n}(\delta) \end{pmatrix}.$$

$H_{2n}(\delta)$ has rows $\{\tilde{\psi}^{\lambda\lambda}_{1n}(\delta), \tilde{\psi}^{\lambda\rho}_{1n}(\delta), \tilde{\psi}^{\rho\lambda}_{1n}(\delta), \tilde{\psi}^{\rho\rho}_{1n}(\delta)\}$ and $\{\tilde{\psi}^{\lambda\lambda}_{2n}(\delta), \tilde{\psi}^{\lambda\rho}_{2n}(\delta), \tilde{\psi}^{\rho\lambda}_{2n}(\delta), \tilde{\psi}^{\rho\rho}_{2n}(\delta)\}$, where

$$\begin{aligned}
&\tilde{\psi}^{\lambda\lambda}_{1n}(\delta) = -2T_{2n}(\lambda) - 6R_{1n}(\delta)R_{2n}(\delta) + 8R^3_{1n}(\delta),\\
&\tilde{\psi}^{\lambda\rho}_{1n}(\delta) = 2Q^{\ddagger}_{1n}(\delta) - 8R_{1n}(\delta)Q^{\dagger}_{1n}(\delta) - 2R_{2n}(\delta)S_{1n}(\delta) + 8R^2_{1n}(\delta)S_{1n}(\delta),\\
&\tilde{\psi}^{\rho\rho}_{1n}(\delta) = -2Q^{\dagger}_{2n}(\delta) - 8S_{1n}(\delta)Q^{\dagger}_{1n}(\delta) + 2R_{1n}(\delta)S_{2n}(\delta) + 8R_{1n}(\delta)S^2_{1n}(\delta),\\
&\tilde{\psi}^{\rho\rho}_{2n}(\delta) = -2K_{2n}(\rho) + S_{3n}(\delta) + 6S_{1n}(\delta)S_{2n}(\delta) + 8S^3_{1n}(\delta)\\
&\tilde{\psi}^{\lambda\rho}_{1n}(\delta) = \tilde{\psi}^{\rho\lambda}_{1n}(\delta) = \tilde{\psi}^{\lambda\lambda}_{2n}(\delta) \text{ and } \tilde{\psi}^{\rho\rho}_{1n}(\delta) = \tilde{\psi}^{\lambda\rho}_{2n}(\delta) = \tilde{\psi}^{\rho\lambda}_{2n}(\delta)
\end{aligned}$$

$H_{3n}(\delta)$ is obtained by taking partial derivatives w.r.t. $\delta'$ for every element of $H_{2n}(\delta)$. It has elements:

$$\begin{aligned}
\tilde{\psi}^{\lambda\lambda\lambda}_{1n}(\delta) &= -6T_{3n}(\lambda) + 6R^2_{2n}(\delta) - 48R^2_{1n}(\delta)R_{2n}(\delta) + 48R^4_{1n}(\delta),\\
\tilde{\psi}^{\lambda\lambda\rho}_{1n}(\delta) &= 12R_{2n}(\delta)Q^{\dagger}_{1n}(\delta) + 12R_{1n}(\delta)Q^{\ddagger}_{1n}(\delta) - 24R_{1n}(\delta)R_{2n}(\delta)S_{1n}(\delta) - 48R^2_{1n}(\delta)Q^{\dagger}_{1n}(\delta)\\
&\quad + 48R^3_{1n}(\delta)S_{1n}(\delta),\\
\tilde{\psi}^{\lambda\rho\rho}_{1n}(\delta) &= 2Q^{\ddagger}_{2n}(\delta) + 16Q^{\dagger 2}_{1n}(\delta) + 8S_{1n}(\delta)Q^{\ddagger}_{1n}(\delta) - 8R_{1n}(\delta)Q^{\dagger}_{2n}(\delta) - 64R_{1n}(\delta)S_{1n}(\delta)Q^{\dagger}_{1n}(\delta)\\
&\quad - 2R_{2n}(\delta)S_{2n}(\delta) - 8R_{2n}(\delta)S^2_{1n}(\delta) + 8R^2_{1n}(\delta)S_{2n}(\delta) + 48R^2_{1n}(\delta)S^2_{1n}(\delta),
\end{aligned}$$

$$\tilde{\psi}_{1n}^{\rho\rho\rho}(\delta) = -2Q_{3n}^{\dagger}(\delta) - 12S_{2n}(\delta)Q_{1n}^{\dagger}(\delta) - 12S_{1n}(\delta)Q_{2n}^{\dagger}(\delta) - 48S_{1n}^{2}(\delta)Q_{1n}^{\dagger}(\delta)$$
$$\quad +24R_{1n}(\delta)S_{1n}(\delta)S_{2n}(\delta) + 2R_{1n}(\delta)S_{3n}(\delta) + 48R_{1n}(\delta)S_{1n}^{3}(\delta),$$
$$\tilde{\psi}_{2n}^{\rho\rho\rho}(\delta) = -6K_{3n}(\rho) + S_{4n}(\delta) + 6S_{2n}^{2}(\delta) + 8S_{1n}(\delta)S_{3n}(\delta) + 48S_{1n}^{2}(\delta)S_{2n}(\delta) + 48S_{1n}^{4}(\delta),$$
$\tilde{\psi}_{1n}^{\lambda\lambda\rho}(\delta) = \tilde{\psi}_{1n}^{\lambda\rho\lambda}(\delta) = \tilde{\psi}_{1n}^{\rho\lambda\lambda}(\delta) = \tilde{\psi}_{2n}^{\lambda\lambda\lambda}(\delta)$, $\tilde{\psi}_{1n}^{\rho\rho\rho}(\delta) = \tilde{\psi}_{2n}^{\lambda\rho\rho}(\delta) = \tilde{\psi}_{2n}^{\rho\lambda\rho}(\delta) = \tilde{\psi}_{2n}^{\rho\rho\lambda}(\delta)$ and $\tilde{\psi}_{1n}^{\lambda\rho\rho}(\delta) = \tilde{\psi}_{1n}^{\rho\lambda\rho}(\delta) = \tilde{\psi}_{2n}^{\lambda\lambda\rho}(\delta) = \tilde{\psi}_{1n}^{\rho\rho\lambda}(\delta) = \tilde{\psi}_{2n}^{\lambda\rho\lambda}(\delta) = \tilde{\psi}_{2n}^{\rho\lambda\lambda}(\delta)$.

**Bootstrap estimates of biases.** The $R$-ratios, $S$-ratios and $Q$-ratios at $\delta = \delta_0$ defined above can all be written as functions of $\theta_0$ and $e_n = \sigma_0^{-1}\epsilon_n$, given $X_n$ and $W_{rN}, r = 1, 2$ and using the relations $M_nB_nX_n = 0$ and $W_{1n}Y_n = G_{1n}\left(X_n\beta_0 + B_n^{-1}\epsilon_n\right)$:

$$\begin{aligned}
R_{1n}(\theta_0, e_n) &= \frac{e_n' M_n(\mu_n + \bar{B}_n e_n)}{e_n' M_n e_n}, \\
R_{2n}(\theta_0, e_n) &= \frac{(\mu_n + \bar{B}_n e_n)' M_n(\mu_n + \bar{B}_n e_n)}{e_n' M_n e_n}, \\
S_{rn}(\theta_0, e_n) &= \frac{e_n' M_n D_{rn} M_n e_n}{e_n' M_n e_n}, \ r = 1, 2, 3, 4, \\
Q_{rn}^{\dagger}(\theta_0, e_n) &= \frac{e_n' M_n D_{rn} M_n(\mu_n + \bar{B}_n e_n)}{e_n' M_n e_n}, \ r = 1, 2, 3, \\
Q_{rn}^{\ddagger}(\theta_0, e_n) &= \frac{(\mu_n + \bar{B}_n e_n)' M_n D_{rn} M_n(\mu_n + \bar{B}_n e_n)}{e_n' M_n e_n}, \ r = 1, 2,
\end{aligned}$$

where $\bar{B}_n = \bar{B}_n(\delta_0)$. As a result, we have $\tilde{\psi}_n = \tilde{\psi}_n(\theta_0, e_n)$ and $H_{rn} = H_{rn}(\theta_0, e_n)$ $r = 1, 2, 3$. The bias terms, $b_{-1}$ and $b_{-3/2}$, can be easily estimated using the general bootstrap procedure for a vector nonlinear parameters, the `Bootstrap Algorithm 2` (BA-2), described in Section 2.

Let $\hat{\delta}_n^{\text{bc2}} = (\hat{\lambda}_n^{\text{bc2}}, \hat{\rho}_n^{\text{bc2}})'$ be the 2nd-order bias-corrected version of $\hat{\delta}_n$. Let $\hat{\beta}_n^{\text{bc}} = \tilde{\beta}(\hat{\delta}_n^{\text{bc2}})$ and $\hat{\sigma}_n^{2,\text{bc}} = \tilde{\sigma}_n^2(\hat{\delta}_n^{\text{bc2}})$. As expected, which can also be inferred from the results given in Section 5.4, the QMLEs can be severely biased and a 2nd-order bias-correction effectively eliminates the bias. To conserve space, we do not report the Monte Carlo results for the finite sample biases of the QMLEs and the bias-corrected QMLEs of the SARAR model.

## 5.3 Improved inferences for regression coefficients

Improved $t$-statistics $t_{\texttt{SARAR}}^{\text{bc}}$ and $t_{\texttt{SARAR}}^{\text{bc2}}$ can be constructed as for the SED or SLD model. Replacing $\hat{\delta}_n$ by $\hat{\delta}_n^{\text{bc2}}$ in the definition of $t_{\texttt{SARAR}}$, we obtain a statistic which is expected to have a better finite sample performance:

$$t_{\texttt{SARAR}}^{\text{bc}} = \frac{c_0'\hat{\beta}_n^{\text{bc}} - c_0'\beta_0}{\sqrt{c_0'\hat{V}_{n1}^{\text{bc}}c_0}}, \tag{37}$$

where $\hat{V}_{n1}^{\text{bc}}$ is $V_{n1}$ evaluated at $\hat{\delta}_n^{\text{bc2}}$, $\tilde{\beta}_n^{\text{bc}}$, $\hat{\sigma}_n^{2,\text{bc}}$, $\hat{\gamma}_n^{\text{bc}}$, and $\hat{\kappa}_n^{\text{bc}}$. The last two are the estimates of $\gamma$ and $\kappa$, the skewness and excess kurtosis of $\epsilon_{n,i}$ involved in $\Gamma_n$.

In order to further improve $t^{\text{bc}}_{\texttt{SARAR}}$, note that given $\tilde{\beta}_n = \tilde{\beta}_n(\delta_0)$, let $\tilde{\beta}^{(r)}_n$ be the $r$th derivative with respect to $\delta_0'$, $r = 1, 2$. Also define $F_n(\rho) = [X_n' B_n'(\rho) B_n(\rho) X_n]^{-1} X_n' B_n'(\rho) B_n(\rho)$ where we have $\tilde{\beta}_n(\delta) = F_n(\rho) A_n(\lambda) Y_n$ and $F^{(r)}_n = F^{(r)}_n(\rho_0)$ is the $r$th derivative with respect to $\delta_0'$, $r = 1, 2$. Assuming that $\mathrm{E}(\tilde{\beta}^{(r)}_n)$ exists and that $\tilde{\beta}^{(r)}_n - \mathrm{E}(\tilde{\beta}^{(r)}_n) = O_p(n^{-1/2})$, $r = 1, 2$, by a Taylor expansion, we have,

$$\begin{aligned} \tilde{\beta}_n(\hat{\delta}_n) - \beta_0 &= \tilde{\beta}_n - \beta_0 + \tilde{\beta}^{(1)}_n(\hat{\delta}_n - \delta_0) + \tfrac{1}{2}\tilde{\beta}^{(2)}_n[(\hat{\delta}_n - \delta_0) \otimes (\hat{\delta}_n - \delta_0)] + O_p(n^{-3/2}), \\ &= b_{0n} + \mathrm{E}(\tilde{\beta}^{(1)}_n)(a_{-1/2} + a_{-1}) + b_{1n} a_{-1/2} + \tfrac{1}{2}\mathrm{E}(\tilde{\beta}^{(2)}_n)(a_{-1/2} \otimes a_{-1/2}) + O_p(n^{-3/2}), \end{aligned}$$

where $b_{0n} = F_n B_n^{-1}\epsilon_n$, $b_{1n} = (-F_n G_{1n} B_n^{-1}\epsilon_n,\ F^{(1)}_n B_n^{-1}\epsilon_n)$, $\mathrm{E}(\tilde{\beta}^{(1)}_n) = (-F_n G_{1n} X_n \beta_0,\ F^{(1)}_n X_n \beta_0)$ and $\mathrm{E}(\tilde{\beta}^{(2)}_n) = (0_{k\times 1},\ -F^{(1)}_n G_{1n} X_n \beta_0,\ -F^{(1)}_n G_{1n} X_n \beta_0,\ F^{(2)}_n X_n \beta_0)$. The expressions for $F^{(1)}_n$ and $F^{(2)}_n$ are given in Appendix A. This leads to a second order expansion for $\mathrm{Var}(\hat{\beta}_n)$ or $\mathrm{Var}(\hat{\beta}^{bc2}_n)$:

$$\mathrm{Var}(\hat{\beta}^{\text{bc2}}_n) = \mathrm{Var}[b_{0n} + \mathrm{E}(\tilde{\beta}^{(1)}_n)(a_{-1/2} + a_{-1}) + b_{1n} a_{-1/2} + \tfrac{1}{2}\mathrm{E}(\tilde{\beta}^{(2)}_n)(a_{-1/2} \otimes a_{-1/2})] + O_p(n^{-2}),$$

where $a_{-1/2} = \Omega_n \tilde{\psi}_n$ and $a_{-1} = \Omega_n H^{\circ}_{1n} a_{-1/2} + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2} \otimes a_{-1/2}) = \Omega_n \tilde{\psi}_n + \Omega_n H_{1n} \Omega_n \tilde{\psi}_n + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})[(\Omega_n \tilde{\psi}_n) \otimes (\Omega_n \tilde{\psi}_n)] = \Omega_n \tilde{\psi}_n + \Omega_n(\tilde{\psi}_n' \otimes H_{1n})\mathrm{vec}(\Omega_n) + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(\Omega_n \otimes \Omega_n)(\tilde{\psi}_n \otimes \tilde{\psi}_n)$, (see Yang, 2015b). As for the two simpler models, one can easily obtain the 2nd-order bias-corrected estimators $\hat{\beta}^{\text{bc2}}_n$ and $\hat{\sigma}^{2,\text{bc2}}_n$, but again Monte Carlo results (not reported for brevity) show that they do not differ much from the corresponding 'plug-in' estimators. A similar two stage bootstrap procedure as given in Section 3, but based on the algorithm `BA-2` presented in Section 2, can be applied to obtain an estimate of this variance term, $\widehat{\mathrm{Var}}(\hat{\beta}^{\text{bc2}}_n)$. We have a second order bias corrected $t$-statistic as follows:

$$t^{\text{bc2}}_{\texttt{SARAR}} = \frac{c_0'\hat{\beta}^{\text{bc2}}_n - c_0'\beta_0}{\sqrt{c_0'\widehat{\mathrm{Var}}(\hat{\beta}^{\text{bc2}}_n)c_0}}. \tag{38}$$

### 5.4 Monte Carlo results

The methods for bias-correction and for improved inferences introduced above for the SARAR model are investigated for their finite sample performance under the following DGP:

$$Y_n = \lambda W_{1n} Y_n + \iota_n \beta_0 + X_{1n}\beta_1 + X_{2n}\beta_2 + u_n, \quad u_n = \rho W_{2n} u_n + \epsilon_n,$$

where all the quantities are generated in a similar manner to those for the SED model. The two spatial weight matrices are taken to be the same. The parameters are set to be the same as before, where $\lambda$ and $\rho$ both take values from $\{-0.5, -0.25, 0, 0.25, 0.5\}$.

We focus on the finite sample performance of the three tests $t_{\texttt{SARAR}}$, $t^{\text{bc}}_{\texttt{SARAR}}$ and $t^{\text{bc2}}_{\texttt{SARAR}}$. The results for the finite sample bias of the QMLEs are available from the authors upon request. Tables 5.1-5.3 report empirical sizes of $t_{\texttt{SARAR}}$, $t^{\text{bc}}_{\texttt{SARAR}}$ and $t^{\text{bc2}}_{\texttt{SARAR}}$ when used for testing $H_0 : \beta_1 = \beta_2$, under the `Group Interaction` spatial layouts described in Appendix B. Similar conclusions are drawn from the Monte Carlo results for the SARAR model as those for the two sub models considered in the earlier sections: (i) as $n$ increases, all tests converge in terms of sizes, (ii) the asymptotic test $t_{\texttt{SARAR}}$ remains unreliable in the sense it rejects the true $H_0$ much too often than it supposes to, (iii) the test $t^{\text{bc}}_{\texttt{SARAR}}$ offers immediate reduction in size distortions, and (iv) $t^{\text{bc2}}_{\texttt{SARAR}}$ generally offers further improvements. Furthermore, like the asymptotic test for the SED model, $t_{\texttt{SARAR}}$ can have a size distortion that is very persistent, having values that are at least 24% even when $n = 500$. The results (unreported to conserve space) under `Rook` and `Queen Contiguity` show similar patterns, but the differences are of a lesser degree due to the weaker spatial dependence (less number of neighbours) under these two spatial layouts.

## 6. Conclusions

This paper considers inference problems for the regression coefficients $\beta$ in linear regression models with spatial dependence, where the estimation of the spatial parameters may incur severe bias. It is shown that while the existence of spatial dependence does not have a big impact on the point estimation of the regression coefficients in terms of consistency and bias (in particular after bias-correcting the spatial estimators), it can have a huge impact on the usual $t$-statistics for $\beta$. We propose simple ways to correct the $t$-statistics, and the resulted 2nd-order corrected $t$-statistics perform superbly. Considering the effectiveness and the simplicity of the proposed methods, they are recommended for practical applications.

Central to the proposed inference methods for regression coefficients in this paper is the general bias-correction methods for nonlinear estimators proposed in Yang (2015b). Thus, the proposed methods have a great potential to be extended to more advanced models such as higher-order SARAR models, spatial panel data models, dynamic panel data models, nonlinear spatial regression models and nonlinear spatial panel data models. They are equally applicable to non-spatial models as well. Among these, the extension to a higher-order SARAR incurs only some extra algebra, and all methods go through in a straightforward manner.

The classical approach to the problem considered in this paper is to directly bootstrap the original $t$-statistic to give asymptotically refined approximations to the finite sample critical values, taking advantage of the underlining statistic being asymptotically pivotal. However, bootstrapping a Wald-type or a likelihood ratio statistic requires the reestimation of all parameters in every bootstrap iteration, and thus is computationally much more demanding compared to our approach, in particular when the model contains more nonlinear parameters that needed to be estimated through numerical optimization (see Yang 2015a for some related works and discussions). Nevertheless, it would be interesting as a future research to compare the two approaches, considering the fact that the direct approach is algebraically simpler.[4]

[4]We thank a referee for raising this issue.

# Appendix A: Additional Quantities for Bias Corrections

**For the SED model**, the full expressions for $D_{jn}(\rho), j = 2, 3, 4$, required in the expressions of $R_{jn}(\rho)$ in (15), for up to third-order bias corrections are:

$$\begin{aligned}
D_{2n}(\rho) &= 2G_n(\rho)P_n(\rho)G_n(\rho) + G_n(\rho)P_n(\rho)G'_n(\rho) - G'_n(\rho)M_n(\rho)G_n(\rho), \\
D_{3n}(\rho) &= \dot{D}_{2n}(\rho) + G_n(\rho)P_n(\rho)D_{2n}(\rho) + D_{2n}(\rho)P_n(\rho)G'_n(\rho) \\
&\quad -G'_n(\rho)M_n(\rho)D_{2n}(\rho) - D_{2n}(\rho)M_n(\rho)G_n(\rho), \\
D_{4n}(\rho) &= \dot{D}_{3n}(\rho) + G_n(\rho)P_n(\rho)D_{3n}(\rho) + D_{3n}(\rho)P_n(\rho)G'_n(\rho) \\
&\quad -G'_n(\rho)M_n(\rho)D_{3n}(\rho) - D_{3n}(\rho)M_n(\rho)G_n(\rho),
\end{aligned}$$

where $P_n(\rho) = I_n - M_n(\rho)$ and $\dot{D}_{jn}(\rho) = \frac{d}{d\rho}D_{jn}(\rho), j = 2, 3$. Note that a predictable pattern emerges from $D_{3n}(\rho)$ onwards. Using the fact that $\frac{d}{d\rho}G_n^i = G_n^{i+1}$ for $i = 1, 2, \ldots$, we have,

$$\begin{aligned}
\dot{D}_{2n}(\rho) &= 2G_n^2(\rho)P_n(\rho)G_n(\rho) - 2G_n(\rho)\dot{M}_n(\rho)G_n(\rho) + 2G_n(\rho)P_n(\rho)G_n^2(\rho) \\
&\quad +G_n^2(\rho)P_n(\rho)G'_n(\rho) - G_n(\rho)\dot{M}_n(\rho)G'_n(\rho) + G_n(\rho)P_n(\rho)G_n'^2(\rho) \\
&\quad -G_n'^2(\rho)M_n(\rho)G_n(\rho) - G'_n(\rho)\dot{M}_n(\rho)G_n(\rho) - G'_n(\rho)M_n(\rho)G_n^2(\rho), \\
\dot{D}_{3n}(\rho) &= G_n'^3(\rho)M_n(\rho)G_n(\rho) + 2G_n'^2(\rho)\dot{M}_n(\rho)G_n(\rho) + 2G_n'^2(\rho)M_n(\rho)G_n^2(\rho) \\
&\quad +G'_n(\rho)\ddot{M}_n(\rho)G_n(\rho) + 2G'_n(\rho)\dot{M}_n(\rho)G_n^2(\rho) + G'_n(\rho)M_n(\rho)G_n^3(\rho) \\
&\quad -2G_n^3(\rho)P_n(\rho)G_n(\rho) + 4G_n^2(\rho)\dot{M}_n(\rho)G_n(\rho) - 4G_n^2(\rho)P_n(\rho)G_n^2(\rho) \\
&\quad +2G_n(\rho)\ddot{M}_n(\rho)G_n(\rho) + 4G_n(\rho)\dot{M}_n(\rho)G_n^2(\rho) - 2G_n(\rho)P_n(\rho)G_n^3(\rho) \\
&\quad -G_n^3(\rho)P_n(\rho)G'_n(\rho) + 2G^2(\rho)\dot{M}_n(\rho)G'_n(\rho) - 2G_n^2(\rho)P_n(\rho)G_n'^2(\rho) \\
&\quad +G_n(\rho)\ddot{M}_n(\rho)G'_n(\rho) + 2G_n(\rho)\dot{M}_n(\rho)G_n'^2(\rho) - G_n(\rho)P_n(\rho)G_n'^3(\rho), \\
\dot{M}_n(\rho) &= P_n(\rho)G'_n(\rho)M_n(\rho) + M_n(\rho)G_n(\rho)P_n(\rho), \\
\ddot{M}_n(\rho) &= 2P_n(\rho)G'_n(\rho)P_n(\rho)G'_n(\rho)M_n(\rho) + 2P_n(\rho)G'_n(\rho)M_n(\rho)G_n(\rho)P_n(\rho) \\
&\quad +2M_n(\rho)G_n(\rho)P_n(\rho)G_n(\rho)P_n(\rho) - 2M_n(\rho)G_n(\rho)P_n(\rho)G'_n(\rho)M_n(\rho).
\end{aligned}$$

The expressions for $F_n^{(r)} = \frac{d^r}{d\rho_0^r}F_n(\rho_0), r = 1, 2$ are:

$F_n^{(1)} = F_n B_n^{-1} G_n^s B_n (X_n F_n - I_n)$, where $G_n^s = G_n + G'_n$
$F_n^{(2)} = F_n^{(1)} B_n^{-1} G_n^s B_n (X_n F_n - I_n) + F_n B_n^{-1} (G_n^{s2} - 2G'_n G_n) B_n (X_n F_n - I_n) + F_n B_n^{-1} G_n^s B_n X_n F_n^{(1)}$

**For the SARAR model**, the full expressions for $D_{jn}(\rho), j = 2, 3, 4$ and $F_n^{(r)}, r = 1, 2$ follow a similar pattern as in the quantities for the SED model with the exception that $G_n(\rho)$ in the SED must now be replaced with $G_{2n}(\rho)$.

## Appendix B: Settings of Monte Carlo Experiments

**Spatial Weight Matrix:** We use three different methods for generating the spatial weight matrix $W_n$: (*i*) `Rook contiguity`, (*ii*) `Queen contiguity`, and (*iii*) `Group Interaction`, with details given in Yang (2015b). The degree of spatial dependence specified by layouts (*i*) and (*ii*) are fixed while in (*iii*) it grows with the increase in sample size. This is attained by allowing for the number of groups, $k$, for each sample to be directly related to $n$. We have considered $k = n^{0.5}$ and $k = n^{0.65}$, where $k$ is the number of groups for each $n$ and hence the degree of spatial dependence indicated by the average group size is $m = n/k$. The actual sizes of the groups are generated from a discrete uniform distribution from $.5m$ to $1.5m$.

**Regressors:** The fixed regressors are generated by `REG1:` $\{x_{1i}, x_{2i}\} \overset{iid}{\sim} N(0,1)/\sqrt{2}$ when Rook or Queen contiguity is followed; and according to either `REG1` or `REG2:` $\{x_{1,ir}, x_{2,ir}\} \overset{iid}{\sim} (2z_r + z_{ir})/\sqrt{10}$, where, $(z_r, z_{ir}) \overset{iid}{\sim} N(0,1)$ when group interaction scheme is followed. The `REG2` scheme gives non-iid regressors where the group means of the regressors' values are different, see Lee (2004). Note that both schemes give a signal-to-noise ratio of 1 when $\beta_1 = \beta_2 = \sigma = 1$.

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

## References


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

Anselin, L., 2001. *Spatial Econometrics. A Companion to Theoretical Econometrics, edited by Badi H. Baltagi.* Blackwell Publishing.

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

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

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

Bao, Y., 2013. Finite sample bias of QMLE in spatial autoregressive models. *Econometric Theory* 29 68-88.

Bao, Y., Ullah, A., 2007. Finite sample properties of maximum likelihood estimator in spatial models. *Journal of Econometrics* 137, 396-413.

Jin, F., Lee, L. F., 2013. Cox-type tests for competing spatial autoregressive models with spatial autoregressive disturbances *Regional Science and Urban Economics.* 43 590-616.

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

Liu S. F., Yang Z. L., 2015. Asymptotic Distribution and Finite-Sample Bias Correction of QML Estimators for Spatial Error Dependence Model. *Econometrics* 3, 376-411.

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

Robinson, P. M., Rossi, F., 2014a. Refined tests for spatial correlation. *Econometric Theory,* forthcoming.

Robinson, P. M., Rossi, F., 2014b. Improved Lagrange multiplier tests in spatial autoregressions. *Econometrics Journal* 17, 139-164.

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

Yang, Z. L., 2015a. LM tests of spatial dependence based on bootstrap critical values. *Journal of Econometrics* 185, 33-39.

Yang, Z. L., 2015b. A general method for third order bias and variance corrections on a non-linear estimator. *Journal of Econometrics* 186, 178-200.

**Table 3.1** Empirical Sizes: Two-Sided Tests of $H_0 : \beta_1 = \beta_2$ in SED Model
`Group Interaction, REG2,` $\sigma = 1$`; Test:` $1 = t_{\texttt{SED}}, 2 = t^{\texttt{bc}}_{\texttt{SED}}, 3 = t^{\texttt{bc2}}_{\texttt{SED}}$

| $\rho$ | Test | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | `Normal Errors` | | | `Normal Mixture` | | | `Lognormal` | | | `Normal Errors` | | | `Normal Mixture` | | | `Lognormal` | | |
| | | $n = 50$ | | | | | | | | | $n = 200$ | | | | | | | | |
| .50 | 1 | .232 | .169 | .088 | .239 | .173 | .092 | .237 | .167 | .086 | .128 | .073 | .020 | .139 | .076 | .024 | .125 | .068 | .018 |
| | 2 | .132 | .078 | .029 | .136 | .083 | .032 | .132 | .079 | .029 | .107 | .057 | .014 | .117 | .060 | .016 | .109 | .059 | .014 |
| | 3 | .113 | .066 | .024 | .116 | .072 | .028 | .112 | .068 | .024 | .100 | .053 | .012 | .108 | .057 | .014 | .102 | .052 | .012 |
| .25 | 1 | .252 | .185 | .102 | .254 | .188 | .104 | .255 | .186 | .100 | .153 | .092 | .030 | .138 | .080 | .026 | .141 | .084 | .028 |
| | 2 | .133 | .083 | .038 | .132 | .083 | .037 | .136 | .082 | .035 | .114 | .060 | .018 | .116 | .063 | .017 | .109 | .058 | .015 |
| | 3 | .117 | .076 | .034 | .120 | .075 | .033 | .121 | .073 | .031 | .108 | .056 | .016 | .109 | .057 | .014 | .101 | .053 | .013 |
| .00 | 1 | .259 | .195 | .105 | .259 | .193 | .105 | .265 | .194 | .104 | .152 | .093 | .035 | .157 | .094 | .034 | .153 | .098 | .034 |
| | 2 | .134 | .084 | .036 | .136 | .086 | .040 | .138 | .085 | .035 | .110 | .060 | .018 | .114 | .062 | .017 | .116 | .066 | .019 |
| | 3 | .125 | .077 | .034 | .125 | .079 | .037 | .126 | .077 | .033 | .104 | .056 | .017 | .108 | .058 | .016 | .108 | .060 | .017 |
| -.25 | 1 | .270 | .195 | .110 | .263 | .193 | .105 | .267 | .196 | .108 | .161 | .101 | .039 | .166 | .102 | .037 | .160 | .099 | .039 |
| | 2 | .141 | .093 | .047 | .140 | .091 | .042 | .146 | .092 | .040 | .114 | .063 | .020 | .114 | .064 | .018 | .111 | .065 | .020 |
| | 3 | .135 | .088 | .044 | .132 | .086 | .040 | .137 | .085 | .037 | .109 | .060 | .018 | .109 | .060 | .018 | .108 | .062 | .019 |
| -.50 | 1 | .260 | .191 | .102 | .258 | .189 | .098 | .262 | .193 | .103 | .166 | .102 | .037 | .167 | .107 | .038 | .168 | .106 | .039 |
| | 2 | .142 | .096 | .043 | .145 | .094 | .044 | .147 | .096 | .043 | .112 | .062 | .017 | .119 | .064 | .020 | .118 | .066 | .020 |
| | 3 | .136 | .099 | .033 | .139 | .099 | .031 | .141 | .099 | .031 | .109 | .060 | .010 | .104 | .060 | .011 | .102 | .061 | .011 |
| | | $n = 100$ | | | | | | | | | $n = 500$ | | | | | | | | |
| .50 | 1 | .164 | .103 | .042 | .170 | .107 | .042 | .172 | .106 | .041 | .123 | .065 | .018 | .124 | .067 | .017 | .120 | .066 | .017 |
| | 2 | .124 | .070 | .023 | .128 | .074 | .021 | .129 | .072 | .018 | .105 | .054 | .014 | .109 | .055 | .013 | .108 | .055 | .013 |
| | 3 | .113 | .062 | .019 | .115 | .064 | .017 | .115 | .062 | .015 | .101 | .053 | .012 | .104 | .051 | .012 | .103 | .052 | .010 |
| .25 | 1 | .190 | .126 | .054 | .192 | .127 | .053 | .192 | .126 | .055 | .132 | .074 | .022 | .126 | .070 | .019 | .130 | .072 | .021 |
| | 2 | .128 | .076 | .023 | .127 | .075 | .021 | .130 | .074 | .025 | .107 | .056 | .015 | .104 | .053 | .014 | .107 | .054 | .015 |
| | 3 | .117 | .068 | .020 | .117 | .067 | .019 | .119 | .067 | .020 | .104 | .053 | .010 | .101 | .051 | .010 | .102 | .052 | .011 |
| .00 | 1 | .200 | .133 | .058 | .197 | .128 | .056 | .204 | .133 | .058 | .132 | .077 | .024 | .136 | .077 | .024 | .134 | .075 | .024 |
| | 2 | .124 | .070 | .024 | .123 | .072 | .023 | .126 | .073 | .025 | .105 | .057 | .015 | .107 | .056 | .014 | .107 | .056 | .015 |
| | 3 | .116 | .064 | .021 | .114 | .066 | .021 | .119 | .070 | .023 | .103 | .050 | .011 | .105 | .051 | .010 | .103 | .051 | .010 |
| -.25 | 1 | .201 | .132 | .060 | .204 | .137 | .059 | .199 | .129 | .057 | .135 | .077 | .023 | .135 | .076 | .021 | .133 | .076 | .021 |
| | 2 | .124 | .072 | .027 | .123 | .071 | .024 | .117 | .068 | .023 | .104 | .056 | .014 | .104 | .053 | .013 | .105 | .056 | .013 |
| | 3 | .116 | .067 | .026 | .115 | .066 | .022 | .109 | .063 | .022 | .102 | .051 | .011 | .102 | .050 | .012 | .102 | .054 | .013 |
| -.50 | 1 | .198 | .137 | .058 | .195 | .130 | .057 | .203 | .133 | .058 | .137 | .077 | .023 | .134 | .077 | .024 | .133 | .077 | .024 |
| | 2 | .118 | .068 | .024 | .117 | .069 | .026 | .120 | .071 | .025 | .105 | .058 | .014 | .100 | .054 | .015 | .101 | .055 | .013 |
| | 3 | .110 | .065 | .021 | .111 | .060 | .020 | .115 | .068 | .021 | .104 | .051 | .010 | .100 | .051 | .011 | .101 | .051 | .010 |

**Table 4.1** Empirical Sizes: Two-Sided Tests of $H_0 : \beta_1 = \beta_2$ in SLD Model

`Group Interaction, REG2,` $\sigma = 1$`; Test:` $1 = t_{\mathtt{SLD}}, 2 = t^{\mathtt{bc}}_{\mathtt{SLD}}, 3 = t^{\mathtt{bc2}}_{\mathtt{SLD}}$

| $\rho$ | Test | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | `Normal Errors` | | | `Normal Mixture` | | | `Lognormal` | | | `Normal Errors` | | | `Normal Mixture` | | | `Lognormal` | | |
| | | $n = 50$ | | | | | | | | | $n = 200$ | | | | | | | | |
| .50 | 1 | .161 | .095 | .028 | .162 | .095 | .029 | .174 | .120 | .058 | .109 | .057 | .013 | .113 | .059 | .014 | .125 | .068 | .019 |
| | 2 | .113 | .062 | .016 | .117 | .068 | .017 | .142 | .098 | .049 | .102 | .051 | .012 | .104 | .052 | .011 | .114 | .061 | .016 |
| | 3 | .095 | .045 | .010 | .100 | .045 | .010 | .100 | .054 | .014 | .093 | .048 | .011 | .098 | .048 | .011 | .100 | .050 | .010 |
| .25 | 1 | .156 | .095 | .028 | .160 | .095 | .026 | .171 | .107 | .041 | .115 | .060 | .013 | .111 | .056 | .012 | .123 | .063 | .014 |
| | 2 | .113 | .062 | .014 | .117 | .065 | .015 | .139 | .085 | .033 | .106 | .055 | .012 | .103 | .050 | .011 | .112 | .054 | .012 |
| | 3 | .095 | .044 | .009 | .093 | .045 | .010 | .098 | .050 | .012 | .100 | .049 | .010 | .097 | .046 | .010 | .100 | .047 | .010 |
| .00 | 1 | .158 | .097 | .030 | .157 | .090 | .028 | .139 | .073 | .021 | .115 | .056 | .012 | .115 | .061 | .014 | .114 | .061 | .015 |
| | 2 | .115 | .065 | .018 | .116 | .062 | .017 | .110 | .060 | .015 | .106 | .051 | .010 | .104 | .053 | .011 | .105 | .054 | .012 |
| | 3 | .100 | .048 | .012 | .093 | .049 | .012 | .099 | .053 | .015 | .099 | .046 | .009 | .099 | .048 | .010 | .098 | .049 | .010 |
| -.25 | 1 | .163 | .096 | .033 | .161 | .099 | .032 | .122 | .067 | .019 | .112 | .057 | .015 | .112 | .058 | .012 | .111 | .058 | .011 |
| | 2 | .117 | .068 | .020 | .124 | .069 | .020 | .100 | .049 | .012 | .105 | .054 | .013 | .105 | .053 | .010 | .108 | .055 | .011 |
| | 3 | .095 | .052 | .014 | .102 | .053 | .015 | .100 | .050 | .010 | .100 | .049 | .012 | .100 | .050 | .010 | .103 | .052 | .009 |
| -.50 | 1 | .167 | .100 | .033 | .161 | .099 | .034 | .113 | .062 | .017 | .119 | .065 | .016 | .108 | .056 | .012 | .105 | .057 | .014 |
| | 2 | .124 | .069 | .020 | .126 | .074 | .022 | .094 | .046 | .012 | .108 | .058 | .013 | .107 | .056 | .013 | .099 | .052 | .011 |
| | 3 | .099 | .055 | .016 | .106 | .051 | .015 | .103 | .050 | .011 | .099 | .051 | .011 | .099 | .049 | .011 | .100 | .051 | .011 |
| | | $n = 100$ | | | | | | | | | $n = 500$ | | | | | | | | |
| .50 | 1 | .131 | .070 | .018 | .127 | .067 | .017 | .133 | .077 | .027 | .106 | .053 | .011 | .111 | .057 | .014 | .107 | .056 | .012 |
| | 2 | .105 | .055 | .013 | .103 | .051 | .011 | .117 | .066 | .023 | .101 | .048 | .009 | .104 | .051 | .012 | .102 | .052 | .011 |
| | 3 | .098 | .054 | .010 | .099 | .049 | .018 | .093 | .048 | .010 | .099 | .050 | .010 | .098 | .048 | .011 | .097 | .048 | .010 |
| .25 | 1 | .127 | .068 | .019 | .130 | .073 | .019 | .145 | .087 | .024 | .110 | .060 | .012 | .109 | .057 | .013 | .109 | .054 | .011 |
| | 2 | .103 | .052 | .014 | .109 | .056 | .014 | .120 | .069 | .018 | .103 | .055 | .011 | .100 | .051 | .010 | .104 | .050 | .010 |
| | 3 | .096 | .049 | .010 | .093 | .050 | .010 | .095 | .046 | .009 | .098 | .050 | .010 | .099 | .050 | .009 | .099 | .049 | .010 |
| .00 | 1 | .133 | .070 | .019 | .130 | .071 | .017 | .128 | .073 | .018 | .107 | .055 | .012 | .109 | .058 | .012 | .108 | .055 | .013 |
| | 2 | .105 | .054 | .013 | .109 | .057 | .012 | .111 | .059 | .013 | .101 | .051 | .011 | .102 | .053 | .010 | .101 | .051 | .012 |
| | 3 | .100 | .050 | .010 | .099 | .050 | .009 | .099 | .052 | .011 | .099 | .050 | .010 | .097 | .049 | .009 | .100 | .049 | .011 |
| -.25 | 1 | .133 | .071 | .020 | .134 | .077 | .021 | .130 | .065 | .013 | .103 | .052 | .014 | .107 | .054 | .012 | .105 | .054 | .012 |
| | 2 | .109 | .054 | .014 | .112 | .060 | .015 | .110 | .051 | .009 | .097 | .048 | .012 | .100 | .049 | .011 | .099 | .050 | .010 |
| | 3 | .100 | .046 | .011 | .099 | .047 | .010 | .103 | .049 | .010 | .099 | .050 | .011 | .099 | .049 | .010 | .099 | .050 | .010 |
| -.50 | 1 | .128 | .071 | .017 | .132 | .074 | .018 | .113 | .057 | .012 | .105 | .056 | .013 | .108 | .056 | .011 | .107 | .054 | .011 |
| | 2 | .106 | .057 | .013 | .112 | .060 | .012 | .094 | .044 | .008 | .098 | .052 | .011 | .102 | .050 | .009 | .100 | .050 | .009 |
| | 3 | .099 | .050 | .010 | .099 | .052 | .011 | .100 | .049 | .010 | .099 | .050 | .010 | .099 | .049 | .010 | .100 | .049 | .009 |

**Table 5.1** Empirical Sizes: Two-Sided Tests of $H_0: \beta_1 = \beta_2$ in SARAR Model

`Group Interaction, REG2,` $\sigma = 1, \lambda = 0.5$; `Test:` $1 = t_{\mathtt{SARAR}}, 2 = t^{\mathtt{bc}}_{\mathtt{SARAR}}, 3 = t^{\mathtt{bc2}}_{\mathtt{SARAR}}$

| $\rho$ | Test | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Normal Errors | | | Normal Mixture | | | Lognormal | | | Normal Errors | | | Normal Mixture | | | Lognormal | | |
| | | $n = 50$ | | | | | | | | | $n = 200$ | | | | | | | | |
| .50 | 1 | .197 | .115 | .040 | .201 | .122 | .044 | .197 | .122 | .040 | .141 | .078 | .022 | .140 | .082 | .028 | .131 | .078 | .021 |
| | 2 | .120 | .068 | .020 | .123 | .073 | .023 | .146 | .084 | .028 | .113 | .056 | .014 | .117 | .061 | .017 | .116 | .061 | .015 |
| | 3 | .115 | .062 | .017 | .119 | .068 | .023 | .128 | .074 | .024 | .105 | .050 | .011 | .107 | .056 | .016 | .106 | .052 | .013 |
| .25 | 1 | .191 | .109 | .031 | .180 | .110 | .031 | .183 | .109 | .035 | .147 | .085 | .025 | .152 | .089 | .028 | .150 | .083 | .025 |
| | 2 | .118 | .067 | .020 | .116 | .069 | .022 | .120 | .066 | .021 | .108 | .056 | .012 | .112 | .061 | .012 | .111 | .058 | .012 |
| | 3 | .109 | .061 | .016 | .103 | .058 | .019 | .103 | .055 | .017 | .100 | .050 | .009 | .102 | .054 | .011 | .101 | .051 | .010 |
| .00 | 1 | .191 | .110 | .031 | .177 | .099 | .028 | .191 | .114 | .037 | .150 | .089 | .026 | .137 | .075 | .016 | .138 | .084 | .020 |
| | 2 | .111 | .054 | .015 | .100 | .054 | .016 | .117 | .065 | .021 | .104 | .055 | .012 | .116 | .061 | .014 | .124 | .066 | .017 |
| | 3 | .098 | .047 | .012 | .095 | .046 | .013 | .100 | .055 | .018 | .097 | .050 | .010 | .102 | .051 | .010 | .105 | .052 | .012 |
| -.25 | 1 | .173 | .100 | .025 | .170 | .096 | .027 | .184 | .108 | .033 | .158 | .093 | .030 | .131 | .074 | .018 | .120 | .062 | .014 |
| | 2 | .094 | .048 | .011 | .098 | .049 | .016 | .111 | .059 | .020 | .108 | .054 | .013 | .123 | .068 | .019 | .118 | .066 | .017 |
| | 3 | .108 | .048 | .009 | .108 | .051 | .013 | .090 | .047 | .016 | .099 | .049 | .012 | .102 | .055 | .010 | .095 | .054 | .010 |
| -.50 | 1 | .182 | .104 | .030 | .162 | .085 | .023 | .177 | .100 | .034 | .127 | .072 | .020 | .120 | .061 | .013 | .119 | .063 | .013 |
| | 2 | .097 | .049 | .013 | .085 | .043 | .010 | .102 | .059 | .019 | .115 | .066 | .017 | .122 | .063 | .015 | .135 | .074 | .019 |
| | 3 | .100 | .048 | .011 | .091 | .052 | .009 | .092 | .046 | .014 | .105 | .060 | .013 | .095 | .046 | .009 | .091 | .052 | .009 |
| | | $n = 100$ | | | | | | | | | $n = 500$ | | | | | | | | |
| .50 | 1 | .169 | .099 | .027 | .163 | .097 | .029 | .171 | .103 | .031 | .124 | .068 | .018 | .126 | .070 | .017 | .124 | .073 | .018 |
| | 2 | .115 | .058 | .014 | .122 | .064 | .017 | .115 | .059 | .016 | .102 | .053 | .013 | .107 | .053 | .011 | .106 | .056 | .011 |
| | 3 | .101 | .049 | .013 | .107 | .057 | .015 | .105 | .055 | .013 | .098 | .049 | .012 | .100 | .049 | .010 | .100 | .050 | .010 |
| .25 | 1 | .165 | .094 | .029 | .172 | .101 | .028 | .163 | .095 | .030 | .130 | .073 | .023 | .134 | .073 | .020 | .130 | .074 | .018 |
| | 2 | .106 | .054 | .012 | .116 | .056 | .011 | .111 | .056 | .013 | .105 | .056 | .015 | .106 | .057 | .014 | .101 | .053 | .012 |
| | 3 | .095 | .045 | .010 | .101 | .047 | .009 | .098 | .049 | .011 | .099 | .052 | .014 | .100 | .053 | .013 | .099 | .049 | .010 |
| .00 | 1 | .177 | .103 | .031 | .176 | .098 | .032 | .165 | .100 | .035 | .138 | .075 | .021 | .135 | .072 | .019 | .133 | .076 | .020 |
| | 2 | .105 | .054 | .012 | .102 | .052 | .013 | .105 | .056 | .014 | .106 | .055 | .013 | .099 | .052 | .009 | .106 | .054 | .012 |
| | 3 | .093 | .046 | .011 | .099 | .048 | .011 | .095 | .051 | .013 | .103 | .053 | .011 | .099 | .049 | .009 | .101 | .053 | .010 |
| -.25 | 1 | .170 | .100 | .027 | .164 | .095 | .029 | .170 | .098 | .029 | .131 | .074 | .020 | .135 | .077 | .022 | .132 | .075 | .022 |
| | 2 | .096 | .047 | .010 | .097 | .048 | .010 | .102 | .048 | .011 | .101 | .055 | .013 | .102 | .053 | .012 | .098 | .053 | .011 |
| | 3 | .095 | .054 | .011 | .099 | .050 | .009 | .099 | .052 | .010 | .096 | .051 | .012 | .099 | .051 | .011 | .100 | .051 | .010 |
| -.50 | 1 | .158 | .091 | .026 | .151 | .086 | .022 | .145 | .087 | .024 | .128 | .071 | .018 | .144 | .076 | .022 | .129 | .072 | .019 |
| | 2 | .090 | .046 | .010 | .091 | .044 | .010 | .091 | .048 | .011 | .094 | .046 | .011 | .107 | .054 | .014 | .093 | .050 | .011 |
| | 3 | .090 | .054 | .009 | .099 | .047 | .009 | .098 | .052 | .009 | .092 | .045 | .011 | .103 | .051 | .013 | .099 | .050 | .010 |

**Table 5.2** Empirical Sizes: Two-Sided Tests of $H_0: \beta_1 = \beta_2$ in SARAR Model

`Group Interaction, REG2,` $\sigma = 1, \lambda = 0.0$; `Test:` $1 = t_{\texttt{SARAR}}, 2 = t^{\texttt{bc}}_{\texttt{SARAR}}, 3 = t^{\texttt{bc2}}_{\texttt{SARAR}}$

| $\rho$ | Test | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Normal Errors | | | Normal Mixture | | | Lognormal | | | Normal Errors | | | Normal Mixture | | | Lognormal | | |
| | | $n = 50$ | | | | | | | | | $n = 200$ | | | | | | | | |
| .50 | 1 | .186 | .107 | .027 | .196 | .127 | .050 | .188 | .114 | .037 | .134 | .079 | .020 | .132 | .066 | .019 | .135 | .074 | .018 |
| | 2 | .130 | .073 | .020 | .123 | .076 | .027 | .132 | .079 | .027 | .111 | .058 | .014 | .107 | .055 | .013 | .116 | .061 | .013 |
| | 3 | .128 | .068 | .017 | .113 | .071 | .025 | .118 | .069 | .023 | .101 | .052 | .012 | .098 | .049 | .010 | .104 | .053 | .010 |
| .25 | 1 | .199 | .125 | .047 | .187 | .113 | .039 | .204 | .130 | .048 | .140 | .083 | .023 | .143 | .086 | .029 | .153 | .087 | .025 |
| | 2 | .115 | .064 | .021 | .126 | .071 | .018 | .111 | .065 | .020 | .117 | .066 | .017 | .107 | .060 | .016 | .108 | .058 | .015 |
| | 3 | .110 | .061 | .020 | .112 | .061 | .015 | .108 | .062 | .019 | .107 | .060 | .014 | .098 | .053 | .014 | .096 | .050 | .013 |
| .00 | 1 | .184 | .110 | .034 | .184 | .107 | .033 | .203 | .126 | .043 | .157 | .093 | .029 | .155 | .094 | .027 | .153 | .089 | .027 |
| | 2 | .110 | .061 | .017 | .114 | .062 | .020 | .127 | .074 | .022 | .106 | .058 | .015 | .112 | .060 | .014 | .110 | .056 | .014 |
| | 3 | .097 | .054 | .015 | .095 | .054 | .017 | .106 | .059 | .017 | .099 | .053 | .013 | .103 | .054 | .013 | .100 | .052 | .012 |
| -.25 | 1 | .192 | .114 | .039 | .189 | .109 | .036 | .194 | .122 | .039 | .127 | .072 | .016 | .136 | .072 | .019 | .127 | .068 | .016 |
| | 2 | .110 | .059 | .018 | .112 | .063 | .017 | .117 | .067 | .021 | .107 | .061 | .014 | .129 | .069 | .019 | .128 | .072 | .019 |
| | 3 | .095 | .050 | .015 | .095 | .051 | .013 | .099 | .055 | .017 | .099 | .053 | .012 | .111 | .054 | .015 | .092 | .049 | .012 |
| -.50 | 1 | .194 | .114 | .038 | .177 | .100 | .030 | .183 | .115 | .033 | .156 | .095 | .028 | .123 | .067 | .014 | .150 | .090 | .030 |
| | 2 | .105 | .058 | .018 | .102 | .052 | .014 | .112 | .062 | .016 | .106 | .053 | .012 | .127 | .071 | .020 | .105 | .056 | .014 |
| | 3 | .098 | .049 | .014 | .098 | .052 | .011 | .099 | .047 | .012 | .098 | .049 | .011 | .105 | .054 | .012 | .096 | .050 | .013 |
| | | $n = 100$ | | | | | | | | | $n = 500$ | | | | | | | | |
| .50 | 1 | .172 | .105 | .030 | .168 | .096 | .032 | .173 | .099 | .030 | .129 | .074 | .021 | .129 | .068 | .016 | .125 | .072 | .017 |
| | 2 | .122 | .067 | .017 | .122 | .065 | .017 | .110 | .054 | .014 | .109 | .055 | .013 | .107 | .053 | .011 | .107 | .057 | .010 |
| | 3 | .107 | .060 | .014 | .109 | .056 | .014 | .100 | .049 | .012 | .102 | .051 | .012 | .099 | .050 | .009 | .100 | .052 | .010 |
| .25 | 1 | .175 | .102 | .030 | .171 | .101 | .031 | .171 | .110 | .036 | .136 | .077 | .018 | .136 | .077 | .022 | .128 | .075 | .019 |
| | 2 | .113 | .057 | .013 | .108 | .055 | .013 | .115 | .064 | .016 | .106 | .053 | .011 | .103 | .057 | .014 | .102 | .053 | .012 |
| | 3 | .098 | .049 | .011 | .096 | .049 | .010 | .105 | .056 | .014 | .100 | .047 | .010 | .099 | .052 | .012 | .097 | .050 | .011 |
| .00 | 1 | .173 | .103 | .030 | .175 | .103 | .034 | .180 | .107 | .031 | .137 | .079 | .021 | .134 | .081 | .022 | .126 | .077 | .021 |
| | 2 | .098 | .051 | .014 | .105 | .056 | .013 | .110 | .054 | .013 | .111 | .053 | .012 | .105 | .055 | .015 | .103 | .057 | .012 |
| | 3 | .097 | .052 | .013 | .094 | .050 | .011 | .098 | .047 | .011 | .105 | .050 | .012 | .099 | .051 | .014 | .100 | .053 | .012 |
| -.25 | 1 | .180 | .109 | .032 | .159 | .094 | .028 | .165 | .099 | .030 | .136 | .077 | .023 | .142 | .082 | .021 | .136 | .071 | .019 |
| | 2 | .104 | .052 | .011 | .094 | .046 | .012 | .101 | .049 | .011 | .101 | .053 | .011 | .109 | .055 | .011 | .099 | .049 | .010 |
| | 3 | .093 | .046 | .010 | .091 | .054 | .010 | .099 | .055 | .011 | .098 | .050 | .010 | .105 | .053 | .010 | .100 | .050 | .010 |
| -.50 | 1 | .172 | .106 | .029 | .159 | .093 | .026 | .158 | .096 | .025 | .146 | .076 | .020 | .134 | .078 | .017 | .138 | .076 | .021 |
| | 2 | .101 | .048 | .011 | .090 | .045 | .009 | .093 | .048 | .009 | .103 | .050 | .010 | .103 | .053 | .009 | .101 | .051 | .011 |
| | 3 | .096 | .054 | .010 | .098 | .049 | .009 | .096 | .054 | .010 | .100 | .047 | .010 | .100 | .050 | .009 | .100 | .049 | .011 |

**Table 5.3** Empirical Sizes: Two-Sided Tests of $H_0 : \beta_1 = \beta_2$ in SARAR Model

`Group Interaction, REG2,` $\sigma = 1, \lambda = -.25$; `Test:` $1 = t_{\texttt{SARAR}}, 2 = t^{\texttt{bc}}_{\texttt{SARAR}}, 3 = t^{\texttt{bc2}}_{\texttt{SARAR}}$

| $\rho$ | Test | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Normal Errors | | | Normal Mixture | | | Lognormal | | | Normal Errors | | | Normal Mixture | | | Lognormal | | |
| | | $n = 50$ | | | | | | | | | $n = 200$ | | | | | | | | |
| .50 | 1 | .196 | .119 | .045 | .203 | .126 | .047 | .188 | .115 | .045 | .129 | .073 | .020 | .144 | .083 | .022 | .133 | .072 | .016 |
| | 2 | .121 | .070 | .020 | .122 | .076 | .022 | .138 | .085 | .030 | .112 | .063 | .014 | .116 | .060 | .012 | .111 | .059 | .013 |
| | 3 | .114 | .066 | .017 | .117 | .072 | .020 | .122 | .074 | .022 | .104 | .055 | .013 | .106 | .054 | .012 | .093 | .046 | .009 |
| .25 | 1 | .198 | .123 | .042 | .205 | .128 | .043 | .205 | .130 | .054 | .143 | .081 | .025 | .150 | .085 | .022 | .151 | .084 | .025 |
| | 2 | .108 | .059 | .018 | .109 | .057 | .020 | .112 | .066 | .025 | .122 | .062 | .018 | .122 | .065 | .015 | .112 | .055 | .014 |
| | 3 | .103 | .056 | .015 | .104 | .055 | .017 | .109 | .064 | .022 | .110 | .056 | .016 | .106 | .057 | .011 | .100 | .051 | .013 |
| .00 | 1 | .192 | .115 | .037 | .180 | .109 | .038 | .199 | .127 | .051 | .144 | .086 | .023 | .129 | .075 | .016 | .156 | .091 | .028 |
| | 2 | .115 | .065 | .017 | .118 | .065 | .017 | .106 | .062 | .020 | .123 | .066 | .017 | .114 | .065 | .015 | .113 | .056 | .013 |
| | 3 | .103 | .058 | .014 | .101 | .056 | .015 | .104 | .059 | .020 | .110 | .059 | .014 | .100 | .053 | .011 | .103 | .050 | .012 |
| -.25 | 1 | .196 | .114 | .032 | .186 | .108 | .038 | .194 | .115 | .042 | .136 | .075 | .023 | .125 | .065 | .018 | .153 | .090 | .026 |
| | 2 | .107 | .052 | .016 | .109 | .060 | .019 | .114 | .069 | .022 | .123 | .068 | .018 | .120 | .062 | .017 | .106 | .056 | .011 |
| | 3 | .099 | .050 | .013 | .098 | .051 | .015 | .098 | .057 | .017 | .112 | .060 | .015 | .101 | .048 | .012 | .097 | .052 | .011 |
| -.50 | 1 | .188 | .113 | .040 | .188 | .111 | .037 | .186 | .116 | .040 | .120 | .064 | .016 | .114 | .055 | .011 | .150 | .091 | .025 |
| | 2 | .111 | .061 | .018 | .089 | .049 | .014 | .098 | .055 | .015 | .117 | .063 | .015 | .126 | .063 | .015 | .105 | .051 | .012 |
| | 3 | .095 | .051 | .015 | .093 | .055 | .013 | .099 | .051 | .015 | .106 | .055 | .012 | .099 | .045 | .009 | .097 | .049 | .011 |
| | | $n = 100$ | | | | | | | | | $n = 500$ | | | | | | | | |
| .50 | 1 | .175 | .100 | .029 | .171 | .099 | .030 | .167 | .098 | .033 | .132 | .069 | .016 | .131 | .070 | .017 | .133 | .072 | .021 |
| | 2 | .116 | .059 | .016 | .126 | .067 | .018 | .113 | .061 | .014 | .110 | .058 | .012 | .108 | .055 | .010 | .109 | .056 | .012 |
| | 3 | .100 | .051 | .013 | .111 | .057 | .015 | .104 | .055 | .014 | .104 | .053 | .011 | .100 | .048 | .010 | .102 | .052 | .011 |
| .25 | 1 | .179 | .102 | .032 | .172 | .103 | .034 | .170 | .099 | .031 | .132 | .079 | .023 | .125 | .074 | .019 | .138 | .076 | .020 |
| | 2 | .114 | .059 | .014 | .111 | .063 | .015 | .108 | .057 | .014 | .109 | .060 | .014 | .106 | .055 | .012 | .107 | .052 | .012 |
| | 3 | .102 | .052 | .012 | .099 | .053 | .013 | .098 | .052 | .011 | .104 | .056 | .013 | .100 | .051 | .009 | .103 | .051 | .011 |
| .00 | 1 | .176 | .106 | .030 | .178 | .103 | .032 | .158 | .093 | .029 | .135 | .077 | .025 | .129 | .077 | .020 | .128 | .071 | .019 |
| | 2 | .099 | .054 | .012 | .108 | .055 | .012 | .096 | .044 | .011 | .105 | .056 | .015 | .099 | .049 | .012 | .100 | .050 | .011 |
| | 3 | .099 | .055 | .010 | .096 | .048 | .010 | .099 | .045 | .011 | .101 | .053 | .013 | .100 | .049 | .011 | .099 | .050 | .011 |
| -.25 | 1 | .177 | .102 | .031 | .165 | .098 | .031 | .162 | .097 | .029 | .139 | .082 | .026 | .139 | .079 | .022 | .130 | .077 | .020 |
| | 2 | .096 | .048 | .009 | .101 | .050 | .013 | .101 | .053 | .013 | .106 | .059 | .014 | .104 | .053 | .011 | .099 | .050 | .012 |
| | 3 | .099 | .052 | .010 | .099 | .050 | .013 | .100 | .050 | .011 | .101 | .056 | .014 | .099 | .050 | .010 | .100 | .050 | .012 |
| -.50 | 1 | .169 | .102 | .029 | .160 | .100 | .032 | .159 | .100 | .035 | .143 | .085 | .023 | .140 | .084 | .024 | .126 | .074 | .023 |
| | 2 | .098 | .047 | .010 | .095 | .049 | .012 | .099 | .053 | .012 | .107 | .054 | .012 | .111 | .059 | .014 | .098 | .053 | .013 |
| | 3 | .096 | .051 | .009 | .096 | .049 | .011 | .099 | .052 | .012 | .105 | .053 | .011 | .108 | .055 | .012 | .099 | .052 | .012 |