1 Introduction

Positive spatial dependence is predominant in the spatial data analyses. Therefore, it is not surprising that most of the methodological papers are concerned with the positive spatial dependence (either in terms of spatial lag, spatial error, or both) when evaluating estimation, testing and forecasting procedures [For example, see Anselin, Bera, Florax, and Yoon (1996) and Baltagi, Song, and Koh (2003) for testing spatial dependence.] However, prevalence of negative spatial dependence is not uncommon as evidenced in many applied papers; just to mention a few: the studies of welfare competition or federal grants competition among local governments [Saavedra (2000) and Boarnet and Glazer (2002)], and the studies of regional employment [Filiztekin (2009) and Pavlyuk (2011)], the cross-border lottery shopping [Garrett and Marsh (2002)], foreign direct investment in OECD countries [Garretsen and Peeters (2009)] and locations of Turkish manufacturing industry [Basdas (2009)]. It appears that negative spatial autocorrelation is likely to occur when competition between regions (or agents) outweigh cooperative factors.

In contrast to the time-series analysis, positive and negative spatial dependence can have quite different implications. Consider a simple first-order autoregressive model,

\[ y_t = \rho y_{t-1} + \epsilon_t, \ |\rho| < 1, \ t = 1, 2, ..., T. \]
where $\epsilon_t \sim iid(0, \sigma^2\epsilon)$, and $y_t \sim (0, \sigma^2)$. The variance-covariance matrix of $y$ has the diagonal elements equal to $\sigma^2/(1 - \rho^2)$ and off-diagonal elements $V_{ij} = \frac{\sigma^2}{1 - \rho^2} |\rho^{i-j}|$. Therefore, the only difference between positive or negative autocorrelation is just the sign of the elements in the matrix. Thus, theoretically there is not much difference between positive or negative autocorrelation in terms of properties of the model. Moreover, previous Monte Carlo studies such as Kramer and Zeisel (1990), King (1985), L’Esperance and Taylor (1975), and Park (1975) suggest that the empirical power functions of various tests for serial autocorrelation, for example Durbin-Watson test and BLUS test, appear to be symmetric around zero and the symmetry becomes more apparent when the sample size $(T)$ grows. In particular, Park (1975) reported the empirical power functions of Durbin-Watson test and Durbin’s h test in the presence of lagged dependent variable where coefficient ($\beta_1$) was fixed at 0.5. His results, shown in Table 1, indicate the symmetry feature of the power for positive and negative values. In his paper, though the experiments were conducted under various values of the coefficient of lagged dependent variable ($\beta_1$), only the empirical power functions under $\beta_1 = 0.5$ was reported. Rayner (1994) repeated the Monte Carlo experiments of Park (1975) and suggests that the results are similar with different values of $\beta_1$. The empirical power functions obtained in Rayner (1994) are shown in Figure 1.

Table 1: Power of the Tests for Serial Autocorrelation, Table 2 of Park (1975)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>-0.95</th>
<th>-0.80</th>
<th>-0.60</th>
<th>-0.40</th>
<th>-0.20</th>
<th>0.00</th>
<th>0.20</th>
<th>0.40</th>
<th>0.60</th>
<th>0.80</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>DW</td>
<td>0.72</td>
<td>0.51</td>
<td>0.23</td>
<td>0.30</td>
<td>0.10</td>
<td>0.00</td>
<td>0.23</td>
<td>0.30</td>
<td>0.57</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>0.89</td>
<td>0.84</td>
<td>0.76</td>
<td>0.34</td>
<td>0.15</td>
<td>0.12</td>
<td>0.21</td>
<td>0.48</td>
<td>0.62</td>
<td>0.82</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Now consider the first-order spatial autoregressive model:

$$Y = \rho W Y + \epsilon,$$

where $Y$ is an $(N \times 1)$ vector of observations, $W$ is an $(N \times N)$ spatial weights matrix and $\epsilon \sim (0, I\sigma^2)$. The variance-covariance matrix of $Y$ can be written as,

$$Var(Y) = (I - \rho W)^{-1}((I - \rho W)')^{-1} \sigma^2\epsilon.$$
Figure 1: Power of the Tests for Serial Autocorrelation, Figures 1 and 2 of Rayner (1994)
Because of the feature of $W$, the structures of the above matrix for positive or negative values of $\rho$ can be very different. For example, under $n = 6$, $\sigma^2 = 1$, $\rho = 0.5$ and $W$ created based on a $3 \times 2$ regular grid with queen criterion, the variance-covariance matrix of $Y$ is

$$
Var(Y) = \begin{bmatrix}
1.406 & 0.602 & 0.319 & 0.671 & 0.602 & 0.319 \\
0.602 & 1.431 & 0.602 & 0.605 & 0.602 \\
0.319 & 0.602 & 1.406 & 0.319 & 0.602 & 0.671 \\
0.671 & 0.602 & 0.319 & 1.406 & 0.602 & 0.319 \\
0.602 & 0.605 & 0.602 & 0.602 & 1.431 & 0.602 \\
0.319 & 0.602 & 0.671 & 0.319 & 0.602 & 1.406
\end{bmatrix},
$$

Under $\rho = -0.5$, the variance-covariance matrix of $Y$ is

$$
Var(Y) = \begin{bmatrix}
1.174 & -0.230 & 0.087 & -0.266 & -0.230 & 0.087 \\
-0.230 & 1.162 & -0.230 & -0.230 & -0.072 & -0.230 \\
0.087 & -0.230 & 1.174 & 0.087 & -0.230 & -0.266 \\
-0.266 & -0.230 & 0.087 & 1.174 & -0.230 & 0.087 \\
-0.230 & -0.072 & -0.230 & -0.230 & 1.162 & -0.230 \\
0.087 & -0.230 & -0.266 & 0.087 & -0.230 & 1.174
\end{bmatrix},
$$

Comparing the two variance-covariance matrices, both the signs and magnitudes of the elements of the two matrices are different. This motivates the question of studying other properties with negative value of spatial autocorrelation.

Previous studies on testing spatial models usually consider only positive spatial dependence. Anselin and Rey (1991) conduct Monte Carlo studies to compare the properties of Moran’s I and Rao score test for spatial dependence, for both spatial error autocorrelation and spatially lagged dependent variable. The empirical power functions of both tests are shown in Figure 2. Though not symmetric around zero, we still see that the power of both tests increase as the true value of spatial autocorrelation coefficient move away from zero. Besides, the asymmetry of the tests becomes less obvious as sample size increases. However, Anselin and Rey (1991) consider spatial dependence in errors and lagged dependent variable separately, not jointly. Anselin et al. (1996) further include
the Monte Carlo studies considering both kinds of spatial dependence and the properties of adjusted Rao score tests, but as mentioned in their study, negative parameter values are excluded to avoid complications.

This paper is concerned with the case of negative spatial dependence and its consequence on estimation, specification tests and calculation of impact effects. We will investigate how negative spatial dependence has bearings upon econometric analysis and in particular, first we will extend the theory and Monte Carlo results in previous literature by including negative coefficients. We will also extend the theoretical derivation and simulations to compare various tests for spatial autocorrelation in Anselin et al. (1996). Then we will specifically

Figure 2: Empirical Power Functions in Anselin and Rey (1991)

(a) Power of Moran’s I

(b) Power of Rao score Test
show how we need to alter the standard methodologies for model specification, estimation, evaluation and forecasting in the presence of negative spatial dependence.

2 A General Approach to testing in the presence of a nuisance parameter

Consider a general statistical model represented by the log-likelihood function $L(\gamma, \psi, \phi)$, where $\gamma$ is a parameter vector, and for simplicity $\psi$ and $\phi$ are taken as scalars to conform with the Spatial Autoregressive model. Suppose an investigator sets $\phi = 0$ and tests $H_0 : \psi = \psi_0$ using the log-likelihood function $L_1(\gamma, \psi) = L(\gamma, \psi, \phi_0)$, where $\psi_0$ and $\phi_0$ are known values. The Rao-Score test statistic for testing $H_0$ under $L_1(\gamma, \psi)$ will be denoted by $RS_\psi$. Let us also denote $\theta = (\gamma', \psi', \phi')'$ and $\tilde{\theta} = (\tilde{\gamma}', \psi_0', \phi_0')$, where $\tilde{\gamma}$ is the maximum likelihood estimator (MLE) of $\gamma$ when $\psi = \psi_0$ and $\phi = \phi_0$. The score vector and the information matrix are defined, respectively, as

$$d(\theta) = \frac{\partial L(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial L(\theta)}{\partial \gamma} \\ \frac{\partial L(\theta)}{\partial \psi} \\ \frac{\partial L(\theta)}{\partial \phi} \end{bmatrix}$$

and

$$J(\theta) = -E \left[ \frac{1}{N} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right] = \begin{bmatrix} J_\gamma & J_{\gamma \psi} & J_{\gamma \phi} \\ J_{\gamma \psi} & J_\psi & J_{\psi \phi} \\ J_{\gamma \phi} & J_{\psi \phi} & J_\phi \end{bmatrix},$$

If $L_1(\gamma, \psi)$ were the true model, it is well known that under $H_0 : \psi = \psi_0$,

$$RS_\psi = \frac{1}{N} d_\psi(\tilde{\theta})' J_{\psi \gamma}^{-1}(\tilde{\theta}) d_\psi(\tilde{\theta}) \overset{D}{\to} \chi^2_1(0),$$

where $J_{\psi \gamma} = J_\psi(\theta) - J_{\psi \gamma} J_\gamma^{-1} J_{\gamma \psi}$. We use $\overset{D}{\to}$ to denote convergence in distribution. Under $H_1 : \psi = \psi_0 + \xi/\sqrt{N}$,

$$RS_\psi \overset{D}{\to} \chi^2_1(\lambda_1).$$ (1)
where the noncentral parameter \( \lambda_1 = \xi'J_{\psi,\gamma}\xi \). Under the set-up, asymptotically the test will have the correct size and will be locally optimal. Now suppose that the true log-likelihood function is \( L_2(\gamma, \phi) = L(\gamma, \psi_0, \phi) \), so the alternative \( L_1(\gamma, \psi) \) becomes completely misspecified. Using a sequence of local values \( \phi = \phi_0 + \delta/\sqrt{N} \), Davidson and MacKinnon (1987) and Saikkonen (1989) obtained the asymptotic distribution of \( RS_\psi \) under \( L_2(\gamma, \phi) \) as

\[
RS_\psi \xrightarrow{D} \chi^2_1(\lambda_2).
\]

where the non central parameter \( \lambda_2 \) is given by \( \lambda_2 = \delta'J_{\phi,\gamma}^{-1}J_{\psi,\gamma}^{-1}J_{\psi,\gamma}\delta \), with \( J_{\psi,\gamma} = J_{\psi\gamma} - J_{\psi\gamma}J_{\gamma,\gamma}^{-1}J_{\gamma,\gamma} \).

Turning to the case of undermisspecification, let the true model be represented by the log-likelihood \( L(\gamma, \psi, \phi) \). The alternative \( L_1(\gamma, \psi) \) is under-specified with respect to nuisance parameter \( \phi \), leading to the problem of undertesting. Consider the locally departures \( \phi = \phi_0 + \delta/\sqrt{N} \) together with \( \psi = \psi_0 + \xi/\sqrt{N} \), Bera and Yoon (1991) derived the asymptotic distribution of \( RS_\psi \),

\[
RS_\psi \xrightarrow{D} \chi^2_1(\lambda_3).
\]

where

\[
\lambda_3 = (\delta'J_{\phi,\gamma} + \xi'J_{\psi,\gamma})J_{\psi,\gamma}^{-1}(J_{\psi,\gamma}\delta + J_{\psi,\gamma}\xi) = \lambda_1 + \lambda_2 + 2\xi'J_{\phi,\gamma}\delta.
\]

Using the result, we can compare the asymptotic local power of the under-specified test with that of the optimal test. It turns out that the contaminated non central parameter \( \lambda \) may increase or decrease the power depending on the configuration of the term \( \xi'J_{\psi,\gamma}\delta \).

Based on (2), Bera and Yoon (1993) suggested a modification to \( RS_\psi \) so the resulting test is valid in the local presence of \( \phi \). The modified statistic is given by
\[ RS^*_\psi = \frac{1}{N} [d_\psi(\hat{\theta}) - J_{\psi,\gamma}(\hat{\theta}) J_{\psi,\gamma}^{-1}(\hat{\theta}) J_{\phi,\gamma}(\hat{\theta})]' \times [J_{\psi,\gamma}(\hat{\theta}) - J_{\psi,\gamma}(\hat{\theta}) J_{\phi,\gamma}^{-1}(\hat{\theta}) J_{\psi,\gamma}(\hat{\theta})]^{-1} \times [d_\psi(\hat{\theta}) - J_{\psi,\gamma}(\hat{\theta}) J_{\phi,\gamma}^{-1}(\hat{\theta}) J_{\phi,\gamma}(\hat{\theta})]. \] 

The modified test statistic essentially adjusts the mean and variance of the standard \( RS^*_\psi \). Bera and Yoon (1993) proved that under \( \psi = \psi_0 \) and \( \phi = \phi_0 + \delta/\sqrt{N} \), \( RS^*_\psi \) has a central \( \chi^2 \) distribution. They also showed that for local misspecification the adjusted test is asymptotically equivalent to Neyman’s \( C(\alpha) \) test and hence shares the optimality properties of the \( C(\alpha) \) test. That is, under the local alternative \( \psi = \psi_0 + \xi/\sqrt{N} \),

\[ RS^*_\psi \xrightarrow{D} \chi^2_1(\lambda_4). \] 

where \( \lambda_4 = \xi' (J_{\psi,\gamma} - J_{\psi,\phi,\gamma} J_{\phi,\gamma}^{-1} J_{\psi,\gamma}) \xi \).

Similarly, we can also obtain \( RS^*_\phi \) to test \( H_0 : \phi = \phi_0 \) in the presence of local misspecification and derive the noncentral parameters of \( RS^*_\phi \) and \( RS^*_\psi \).

If \( L_2(\gamma, \phi) \) is the true log-likelihood function, under the null hypothesis \( RS^*_\phi \) asymptotically follows central \( \chi^2 \) distribution, and under local alternative \( H_1 : \phi = \phi_0 + \delta/\sqrt{N} \),

\[ RS^*_\phi \xrightarrow{D} \chi^2_1(\lambda_5). \] 

where \( \lambda_5 = \delta' J_{\phi,\gamma} \delta \). In the case of complete misspecification, we have

\[ RS^*_\phi \xrightarrow{D} \chi^2_1(\lambda_6). \] 

where \( \lambda_6 = \xi' J_{\psi,\gamma} J_{\phi,\gamma}^{-1} J_{\psi,\gamma} \xi \). And in the case of undermisspecification,

\[ RS^*_\phi \xrightarrow{D} \chi^2_1(\lambda_7). \] 

where \( \lambda_7 = \lambda_5 + \lambda_6 + 2\delta' J_{\psi,\gamma} \xi \).

On the other hand, the adjusted Rao-Score test statistic for testing \( H_0 : \phi = \phi_0 \) will follow a central \( \chi^2 \) distribution under the null hypothesis even in the presence of locally misspecification of \( \psi \). And under the local alternative \( \phi = \phi_0 + \delta/\sqrt{N} \),
\[ RS^*_\phi \xrightarrow{D} \chi^2_1(\lambda_8). \]  

where \( \lambda_8 = \delta'(J_{\phi,\gamma} - J_{\psi,\gamma}J_{\psi,\gamma}^{-1}J_{\phi,\gamma})\delta. \)

3 Tests for SARMA Model

To make the study comparable to previous literature on spatial analyses, we consider a general model, the mixed regressive spatial autoregressive moving average (SARMA) model, as specified in Anselin et al. (1996):

\[ y = X\gamma + \phi Wy + u, \]
\[ u = \psi W\epsilon + \epsilon, \]  
\[ \epsilon \sim N(0, \sigma^2 I). \]

where \( y \) is an \((N \times 1)\) vector of observations of dependent variable, \( X \) is an \((N \times k)\) matrix of observations of exogenous variables, and \( \gamma \) is a \((k \times 1)\) vector of parameters. \( \phi \) and \( \psi \) are scalar spatial parameters, and \( W \) is a \((N \times N)\) spatial weights matrix.

We are interested in testing \( H_0 : \psi = 0 \) in the presence of the nuisance parameter \( \phi \). Let \( \theta = (\gamma', \psi, \phi)' \), following the result of Anselin (1988a), we have the following equations:

\[ \frac{\partial L}{\partial \gamma} = d_\gamma = \frac{1}{\sigma^2} X' u, \]
\[ \frac{\partial L}{\partial \psi} = d_\psi = \frac{1}{\sigma^2} u' W u, \]
\[ \frac{\partial L}{\partial \phi} = d_\phi = \frac{1}{\sigma^2} u' W y, \]  

and

\[ J = \begin{bmatrix} X'X & 0 & X'(WX\gamma) \\ 0 & T\sigma^2 & T\sigma^2 \\ (WX\gamma)'X & T\sigma^2 & (WX\gamma)'(WX\gamma) + T\sigma^2 \end{bmatrix} \]

where \( T = tr[(W' + W)W] \). The above equations imply that
\[ J_{\psi_{\gamma}} = J_{\psi_{\gamma}} = J_{\phi_{\psi\gamma}} = \frac{T}{N}, \]
\[ J_{\phi_{\gamma}} = \frac{1}{N\sigma^2} [(WX_\gamma)'M(WX_\gamma) + T\sigma^2] \]
\[ = \frac{T}{N} + \frac{1}{N\sigma^2} (WX_\gamma)'M(WX_\gamma). \]

where \( M = I - X(X'X)^{-1}X' \). The adjusted Rao-Score test statistic can be constructed as,
\[ RS^*_\psi = \frac{[\tilde{u}'W\tilde{u}/\tilde{\sigma}^2 - T(N\tilde{J}_{\phi_{\gamma}})^{-1}\tilde{u}'Wy/\tilde{\sigma}^2]^2}{T[1 - T(N\tilde{J}_{\phi_{\gamma}})^{-1}]} \]  

where \( \tilde{u} = y - X\tilde{\gamma} \) are the OLS residuals, and \( \tilde{\sigma}^2 = \tilde{u}'\tilde{u}/N \), and from (13) it follows that
\[ (N\tilde{J}_{\phi_{\gamma}})^{-1} = \tilde{\sigma}^2[(WX_{\tilde{\gamma}})'M(WX_{\tilde{\gamma}}) + T\tilde{\sigma}^2]^{-1}. \]

The conventional one-directional test \( RS_\psi \) given in Burridge (1980) is obtained by setting \( \phi = 0 \) to yield
\[ RS_\psi = \frac{[\tilde{u}'W\tilde{u}/\tilde{\sigma}^2]^2}{T}. \]  

To see the behavior of \( RS_\psi \) and \( RS^*_\psi \), let us consider the case of locally misspecification, i.e. \( \phi = \phi_0 + \delta/\sqrt{N} \). Under the null hypothesis that \( \psi = 0 \), the noncentral parameter of \( RS_\psi \) is
\[ \lambda_2 = \delta'(\frac{T}{N})\delta \]  
while under the alternative that \( \psi = \psi_0 + \xi/\sqrt{N} \), i.e. under the case of undermisspecification, the noncentral parameter of \( RS_\psi \) is
\[ \lambda_3 = \lambda_1 + \lambda_2 + 2\xi'J_{\phi_{\psi\gamma}}\delta \]
\[ = \xi'(\frac{T}{N})\xi + \delta'(\frac{T}{N})\delta + 2\xi'(\frac{T}{N})\delta. \]  

From the above equations, we see that the noncentral parameters of \( RS_\psi \) under both the null and alternative are affected by the locally misspecification of \( \phi \). In particular, the noncentral parameter of \( RS_\psi \) is affected by the combination
of positive or negative values of $\psi$ and $\phi$ under alternative hypothesis because of the interaction term of $\xi$ and $\delta$. This means the power of the test of $RS_\psi$ would be affected by different combination of positive or negative values of $\psi$ and $\phi$. Comparing to the case where both of the spatial autocorrelation parameters are positive, the noncentral parameter is lower when one parameter is positive and the other is negative. The noncentral parameter can be as low as 0 when $\xi = -\delta$.

On the other hand, the noncentral parameter of $RS^*_\psi$ under the alternative is not affected in the presence of locally misspecification of $\phi$. Under the alternative that $\psi = \psi_0 + \xi/\sqrt{N}$, the noncentral parameter of $RS^*_\psi$ is

$$\lambda_4 = \xi'(\frac{T}{\sqrt{N}} - \frac{T}{\sqrt{N}}^2 \hat{J}_\phi^{-1})\xi$$
$$= \xi'(\frac{T}{\sqrt{N}})[1 - \frac{T\sigma^2_\phi}{T\sigma^2_\phi + (WX\gamma)'M(WX\gamma)}]\xi.$$ \hspace{1cm} (18)

From the expression of the above, the noncentral parameter of $RS^*_\psi$ is not affected by the locally misspecification of $\phi$, nor is it affected by positive or negative alternative value of $\psi$.

We can also study the performance of $RS^*_\phi$ and $RS_\phi$ in the presence of locally misspecification, i.e. $\psi = \psi_0 + \xi/\sqrt{N}$. The unadjusted Rao score test statistic for testing $H_0 : \phi = \phi_0$ is

$$RS_\phi = \frac{[\hat{u}'Wy/\hat{\sigma}^2]^2}{NJ_{\phi,\gamma}}$$ \hspace{1cm} (19)

while the adjusted test statistic is constructed as,

$$RS^*_\phi = \frac{[\hat{u}'Wy/\hat{\sigma}^2 - \hat{u}'\hat{u}/\hat{\sigma}^2]^2}{NJ_{\phi,\gamma} - T}$$ \hspace{1cm} (20)

Under the alternative $\phi = \phi_0 + \delta/\sqrt{N}$, the noncentral parameter of $RS^*_\phi$ is

$$\lambda_8 = \delta'(\frac{1}{N\sigma^2})(WX\gamma)'M(WX\gamma)\delta$$ \hspace{1cm} (21)

while the noncentral parameter of $RS_\phi$ is
\[ \lambda_7 = \delta'J_{\psi,\gamma}\delta + \xi'J_{\psi,\gamma}J_{\psi,\gamma}^{-1}J_{\psi,\gamma}\xi + 2\delta J_{\psi,\gamma}\xi \]

\[ = \delta'\left( \frac{1}{N\sigma^2} \right) [(WX\gamma)'M(WX\gamma) + T\sigma^2] \delta \]

\[ + \xi'\left( \frac{T^2}{N^2} \right) [(WX\gamma)'M(WX\gamma) + T\sigma^2]^{-1} \xi + 2\delta'\left( \frac{T}{N} \right) \xi \]  

Again we see that the noncentral parameter of \( RS_{\phi} \) under the alternative in the presence of locally misspecification of \( \psi \) is affected by the combination of positive or negative values of \( \psi \) and \( \phi \), but the noncentral parameter of \( RS_{\phi}^* \) is not.

### 4 Empirical Applications

To gain more insights on how negative spatial dependence would affect model specification tests and estimation in the real context, we first examine the various test statistics and estimated spatial autocorrelation coefficients in previous literature as illustrative examples. Table 2 shows three examples with positive spatial dependence. All of the three examples share the features that (i) the unadjusted one-directional tests are strongly significant and the joint tests are moderately significant, while the adjusted statistics are lower than the unadjusted ones and show less significance, (ii) the spatial coefficients are positive and significant. The examples show that it is likely that both the joint and unadjusted one-directional tests are spurious because of only one source of spatial dependence.

On the other hand, the cases with negative spatial dependence are more complicated. Table 3 summarizes some empirical results with negative spatial coefficients. As indicated in the previous section, the unadjusted test statistics can be higher or lower than the adjusted ones, depending on the combinations of the signs of two sources of spatial dependence. For example, Garret and Marsh (2002) estimated the revenue impact of cross-border lottery sales for 105 counties in Kansas and found negative spatial autocorrelation for both spatial lag and spatial error coefficients. However, it should be noted that in their study they estimated the two coefficients separately, not jointly. According to
the reported values of Rao-score test statistics and the theoretical prediction in previous sections, we expect that both of the coefficients are negative.

Though there are many empirical studies that found negative spatial dependence, most of the studies only estimate one spatial autocorrelation coefficient, either spatial lag or spatial error, and reported one-directional Rao-score test result. Others consider both two kinds of spatial dependence, but estimate the coefficients separately, as we see in Garret and Marsh (2002), and Basdas (2009). Therefore, to further see the empirical applications of the interactions between the two kinds of spatial dependence with negative values, it is necessary to reinvestigate the data that finds negative spatial autocorrelations, which provides a direction of future research.

5 Monte Carlo Simulations

In this section we present the results of a Monte Carlo study to investigate the finite sample behavior of the tests. We focus specifically on the power of the tests and the comparison of adjusted test relative to unadjusted one. All the tests are based on estimation by OLS. To facilitate comparison with existing results we follow a structure similar to the one adopted by Anselin et al. (1996). The model under the null hypothesis of no spatial dependence is the classic regression model:

\[
\begin{align*}
Y_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \\
\epsilon_i &= \phi \epsilon_{i-1} + \psi Z_i + \eta_i
\end{align*}
\]

Table 2: Summary of Empirical Studies - Positive Dependence

<table>
<thead>
<tr>
<th></th>
<th>( RS_{\psi^0} )</th>
<th>( RS_{\psi} )</th>
<th>( RS_{\psi^0}^* )</th>
<th>( RS_{\phi} )</th>
<th>( \phi )</th>
<th>( \psi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anselin (1988)</td>
<td>9.44</td>
<td>5.72</td>
<td>0.88</td>
<td>9.36</td>
<td>3.72</td>
<td>0.431</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.01)</td>
<td>(0.02)</td>
<td>(&lt;0.01)</td>
<td>(0.05)</td>
<td>(&lt;0.01)</td>
<td>(&lt;0.01)</td>
</tr>
<tr>
<td>Florax (1992)</td>
<td>7.97</td>
<td>2.43</td>
<td>0.14</td>
<td>7.83</td>
<td>5.54</td>
<td>0.349</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.12)</td>
<td>(0.02)</td>
<td>(&lt;0.01)</td>
<td>(&lt;0.01)</td>
<td>(&lt;0.01)</td>
</tr>
<tr>
<td>Anselin et al. (1996)</td>
<td>5.07</td>
<td>4.35</td>
<td>3.65</td>
<td>1.42</td>
<td>0.72</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.04)</td>
<td>(0.23)</td>
<td>(0.40)</td>
<td>(0.11)</td>
<td>(&lt;0.01)</td>
</tr>
</tbody>
</table>

*p-values in parentheses.

Table 3: Summary of Empirical Studies - Negative Dependence

<table>
<thead>
<tr>
<th></th>
<th>( RS_{\psi^0} )</th>
<th>( RS_{\psi} )</th>
<th>( RS_{\psi^0}^* )</th>
<th>( RS_{\phi} )</th>
<th>( \phi )</th>
<th>( \psi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garret &amp; Marsh (2002)</td>
<td>3.91</td>
<td>0.57</td>
<td>0.12</td>
<td>3.79</td>
<td>3.34</td>
<td>-0.064*</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.14)</td>
<td>(0.45)</td>
<td>(0.73)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>Pavlyuk (2011)</td>
<td>7.37</td>
<td>0.03</td>
<td>5.00</td>
<td>2.37</td>
<td>7.34</td>
<td>-1.91*</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.86)</td>
<td>(0.03)</td>
<td>(0.12)</td>
<td>(&lt;0.01)</td>
<td></td>
</tr>
<tr>
<td>Basdas (2009)</td>
<td>6.26</td>
<td>3.77</td>
<td>5.43</td>
<td>0.83</td>
<td>2.49</td>
<td>-0.48*</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.02)</td>
<td>(0.36)</td>
<td>(0.11)</td>
<td></td>
</tr>
</tbody>
</table>

*Estimate coefficients separately.
\[ y = X\gamma + u \]

while under the SARMA alternative the model is specified in equation (10). The \( N \) observations on the dependent variables are generated from a vector of standard normal random variables \( u \). The explanatory variables \( X \), an \( N \times 3 \) matrix is obtained from a vector of a constant term combined with two variates drawn from a uniform \((0, 10)\) distribution. The coefficients of explanatory variables (\( \gamma \)) is set to be a vector of ones. The matrix of explanatory variables is held fixed in the replications. For each combination of parameter values, 5,000 replication were carried out. The graphs are based on the theoretical size of 0.05, and the proportion of rejections (i.e. the proportion of times the computed test statistic exceeded its asymptotic value) is reported.

In order to make comparison with Anselin et al. (1996), the configurations used to generate spatial dependence are formally expressed in three spatial weight matrices. These correspond to sample size of 40 and 81. The weight matrices of size 40 is built from actual irregularly shaped regionalizations of the Netherland (see Florax, 1992 for more details.) The weight matrices for \( N = 81 \) correspond to a regular square 9 \( \times \) 9 grid, with contiguity defined by the rook and queen criterion respectively.

Figures 3 to 5 are the power functions of the \( RS^*_\psi \) and \( RS_\psi \) tests. As theoretical derivation predicts, the power functions of \( RS^*_\psi \) are U-shaped (symmetric around zero) under different values of \( \phi \), while the power functions of \( RS_\psi \) do not have symmetric feature and are sensitive to different values of \( \phi \). Also note that the power is extremely low when the two spatial autocorrelation coefficients have different signs but similar magnitude. Similar results can be shown under the specification of spatial weight matrices based on regular grids, which are presented in Figures 4 and 5. The adjusted test behaves well in the sense of the symmetry of the empirical power function except for the case that spatial weight matrices are built based on queen criterion and there is moderate spatial dependency in the dependent variable (i.e. the nuisance parameter, \( |\phi| = 0.5 \)). One possible explanation is that queen criterion impose too much spatial relationship since it counts all of the 8 directions as one’s neighbors, and the
Figure 3: Power of $RS^*_\psi$ and $RS_\psi$, $N = 40$

moderately spatial dependency strengthen the relationship further. However, since we are considering local departure of the parameters, we can still conclude that in the presence of negative spatial dependence, the adjusted Rao-score test performs better than unadjusted one in the sense of the power of the test.

As for the empirical power functions of $RS^*_\phi$ and $RS_\phi$ tests, both of them have a nicely U-shape around zero. The results are all similar under different design of spatial weight matrices, which can be seen in Figure 6. Still the symmetry is more apparent in $RS^*_\phi$ tests than $RS_\phi$, but the discrepancies are not as large as the one-directional test of the hypothesis $H_0: \psi = 0$.

6 Conclusion

This paper extends the theoretical derivation and Monte Carlo studies of model specification tests in spatial regression by examining the effect of negative spatial dependence. Previous studies focus on positive spatial autocorrelations, and
Figure 4: Power of $RS_{\psi}^*$ and $RS_{\psi}$, $N = 81$, $W$ with rook design
Figure 5: Power of $RS_{\psi}^*$ and $RS_{\psi}$, $N = 81$, $W$ with queen design
Figure 6: Power of $RS_\phi^*$ and $RS_\phi$
therefore only address the over-sized problem of Rao-score tests. Our study suggests that under negative spatial dependence, the power of the conventional Rao-score tests can be very low, and hence, cautious is required when negative dependence is expected and the one-directional Rao-score test conclude no spatial dependence for the errors. By deriving the noncentral parameters of the asymptotic distributions of the test statistics, we are able to explain the low power of the unadjusted Rao-score test in some specific cases, and it can be shown that the power is especially low when one of the source of spatial dependence is positive while the other is negative, and they have similar magnitude. Monte Carlo results are consistent with theoretical prediction even when the sample size is finite.

There are some extensions to our study. First, it will shed more light on how negative spatial dependence affect model specification tests and estimation in the real context by looking at the empirical applications. Besides, since all the test statistics and the noncentral parameters include the spatial weight matrices \((W)\), it would be important to look at how the formation of \(W\) would affect the results, especially when the Monte Carlo studies suggest that higher or lower spatial relations induced by different designs of \(W\) may make a difference, the theoretical comparison among different spatial weight matrices worths exploration. Finally, it would pose valuable applications to further study how negative spatial dependence affects the calculation of impact effects and the evaluation of model forecasting.
References


