$DOI:\,10.1515/JSSI\text{-}2015\text{-}0561$ 

# Testing for Spatial Lag Effects in Varying Coefficient Spatial Autoregressive Models

### Shuang GUO

School of Science, Minzu University of China, Beijing 100081, China E-mail: y1991y.me@163.com

## Chuanhua WEI\*

School of Science, Minzu University of China, Beijing 100081, China E-mail: chweisd@163.com

**Abstract** This paper is concerned with testing for the varying coefficient spatial autoregressive models. Based on the profile likelihood estimation procedure, a profile generalized likelihood ratio test procedure is proposed to test spatial lag effects, and a residual-based bootstrap procedure is used to derive the p-value of the test. Some simulations are conducted to assess the performance of the test and the results are satisfactory.

**Keywords** varying coefficient spatial autoregressive model; profile likelihood estimation; profile generalized likelihood ratio test; local linear method; bootstrap

#### 1 Introduction

In the last three decades, spatial autoregressive models as a popular spatial econometric method have received much attention in the literature. For a long time, however, its main theory has centered around parametric models in which the relationship between the response and the covariates is assumed to be linear. In recent years, many useful semiparametric spatial autoregressive models have been proposed to relax traditional parametric models. Su and Jin<sup>[1]</sup> proposed a partially linear spatial autoregressive model and studied the properties of the profile quasi-maximum likelihood estimators for the parameters in the model. Li and Mei<sup>[2]</sup> proposed a statistical test procedure to check a polynomial relationship of the non-parametric component in partially linear spatial autoregressive models. Su<sup>[3]</sup> studied a nonprametric spatial autoregressive model that the spatially lagged response variable enters the model linearly while the covariates enter the model nonparametrically.

Like parametric models, semiparametric models have various forms. An alternative approach to relax the conditions imposed on traditional parametric models and explore the hidden structure is varying coefficient models, which was introduced by Cleveland et al.<sup>[4]</sup> and popularized by Hastie and Tibshirani<sup>[5]</sup>. Varying coefficient model is a useful extension of linear

Received April 12, 2015, accepted July 22, 2015

Supported by Beijing Higher Education Young Elite Teacher Project (YETP1316), and the National Natural Science Foundation of China (11301565)

<sup>\*</sup>The corresponding author

562 GUO S, WEI C H.

regression model, by allowing all the regression coefficients to vary as unknown functions of other factors. Due to its flexibility, varying coefficient model has been studied in many different contexts and has been successfully applied to nonlinear time series analysis, longitudinal and functional data analysis, and time-varying models in finance, see Fan and Zhang<sup>[6]</sup> for a comprehensive survey. In this paper, we consider the following varying coefficient spatial autoregressive model by combining the spatial autoregressive model and the varying coefficient regression model,

$$Y_i = \rho \sum_{j=1}^n w_{ij} Y_j + \boldsymbol{X}_i^{\mathrm{T}} \boldsymbol{\alpha}(U_i) + \varepsilon_i, \quad i = 1, 2, \cdots, n$$
 (1)

where  $Y_i$ 's are responses,  $X_i = (X_{i1}, X_{i2}, \dots, X_{ip})^{\mathrm{T}}$  and  $U_i$  are associated covariates,  $\boldsymbol{\alpha}(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \dots, \alpha_p(\cdot))^{\mathrm{T}}$  is a p-dimensional vector of unknown functions.  $\boldsymbol{W} = (w_{ij}), 1 \leq i, j \leq n$  is a specified  $n \times n$  spatial weight matrix,  $\varepsilon_i$ 's are independent and identically distributed random errors with zero mean and finite variance  $\sigma^2$ . Obviously, when  $\boldsymbol{\alpha}(\cdot) = \boldsymbol{\alpha}$ , model (1.1) becomes the standard spatial autoregressive model, which was studied by Cliff and  $\mathrm{Ord}^{[7]}$ , Anselin<sup>[8]</sup>, Kelejian and Prucha<sup>[9]</sup> and Lee<sup>[10,11]</sup>. When  $\rho = 0$ , the model (1.1) reduces to the varying coefficient model.

For model (1), Li and Chen<sup>[12]</sup> proposed a profile likelihood approach based on the local linear method to estimate the unknown coefficient functions  $\alpha(\cdot)$  and the spatial lag parameter  $\rho$ . As we all know, an important question of spatial autoregressive model is to test the existence of the spatial effects. This leads to the following testing problem

$$H_0: \rho = 0 \quad \text{VS} \quad H_1: \rho \neq 0$$
 (2)

For the standard linear spatial autoregressive models, Likelihood ratio test and Rao's Score (Lagrange Multiplier) test can be applied to solve this problem, details can be found in Anselin<sup>[8]</sup>. However, the test problem (2) is a semiparametric hypothesis versus another semiparametric hypothesis testing problem. Many traditional tests cannot be directly applied to the above hypothesis. For this kind of testing problem, Fan and Huang<sup>[13]</sup> proposed a profile generalized likelihood ratio (PGLR) test based on the partially linear varying coefficient model. Thus, we are motivated to extend the PGLR test procedure for the testing problem (2) of model (1).

The rest of this paper is organized as follows. In Section 2, the profile maximum likelihood method to fit the varying coefficient spatial autoregressive model is briefly described to facilitate our consequent discussions. In Section 3, the test statistic is proposed and the residual-based bootstrap procedure is suggested to derive the p-value of the test. Simulations are conducted in Section 4 to examine the finite sample performance of the proposed test procedure. Conclusion is presented in Section 5.

### 2 Profile Likelihood Estimation Procedure

For the need of constructing the test statistic, we first introduce the profile likelihood estimating approach of model (1) proposed by Li and Chen<sup>[12]</sup>. Let us work with the matrix notation. Denote

$$oldsymbol{Y} = \left[ egin{array}{c} Y_1 \ Y_2 \ dots \ Y_n \end{array} 
ight], \quad oldsymbol{X} = \left[ egin{array}{c} oldsymbol{X}_1^{
m T} \ oldsymbol{X}_2^{
m T} \ dots \ oldsymbol{X}_n^{
m T} \end{array} 
ight], \quad oldsymbol{M} = \left[ egin{array}{c} oldsymbol{X}_1^{
m T} oldsymbol{lpha}(U_1) \ oldsymbol{X}_2^{
m T} oldsymbol{lpha}(U_2) \ dots \ oldsymbol{X}_n^{
m T} oldsymbol{lpha}(U_n) \end{array} 
ight], \quad oldsymbol{arepsilon} = \left[ egin{array}{c} arepsilon_1 \ oldsymbol{arepsilon} \ oldsymbol{arepsilon} \ oldsymbol{lpha} \ oldsymbol{lpha} \ oldsymbol{lpha} \end{array} 
ight]$$

Then model (1) can be written as

$$Y = \rho WY + M + \varepsilon \tag{3}$$

Assume that  $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ , let  $\boldsymbol{\theta} = (\rho, \sigma^2)^{\mathrm{T}}$ , the log-likelihood function of model (3) is

$$\log L_n(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\alpha}(U_1), \boldsymbol{\alpha}(U_2), \cdots, \boldsymbol{\alpha}(U_n))$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) + \log|\boldsymbol{I}_n - \rho \boldsymbol{W}| - \frac{(\boldsymbol{Y} - \rho \boldsymbol{W} \boldsymbol{Y} - \boldsymbol{M})^{\mathrm{T}}(\boldsymbol{Y} - \rho \boldsymbol{W} \boldsymbol{Y} - \boldsymbol{M})}{2\sigma^2}$$
(4)

where  $I_n$  is the identity matrix of order n.

If the parameter  $\rho$  is known, then model (1) can be written as

$$Y_i^* = \alpha_1(U_i)X_{i1} + \alpha_2(U_i)X_{i2} + \dots + \alpha_p(U_i)X_{ip} + \varepsilon_i$$
(5)

where  $(Y_1^*, Y_2^*, \dots, Y_n^*) = \mathbf{Y} - \rho \mathbf{W} \mathbf{Y} = (\mathbf{I}_n - \rho \mathbf{W}) \mathbf{Y}$ . This transforms the varying coefficient spatial autoregressive model (1) into the standard varying coefficient model (5). Now, we apply a local linear regression technique to estimate the varying coefficient functions  $\{\alpha_j(\cdot), j = 1, 2, \dots, p\}$  in the model (5). For u in a small neighborhood of  $u_0$ , one can approximate  $\alpha_j(\cdot)$  locally by a linear function

$$\alpha_j(u) \approx \alpha_j(u_0) + \alpha'_j(u_0)(u - u_0), \quad j = 1, 2, \dots, p$$

This leads to the following weighted local least-squares problems: find  $\left[\alpha_j(u_0), \alpha'_j(u_0)\right]$  to minimize

$$\sum_{i=1}^{n} \left[ Y_i^* - \sum_{j=1}^{p} \{ \alpha_j(u_0) + \alpha_j'(u_0)(U_i - u_0) \} X_{ij} \right]^2 K_h(U_i - u_0)$$
 (6)

where K is a kernel function, h is a bandwidth and  $K_h(\cdot) = K(\cdot/h)/h$ .

Let

$$m{D}_{u_0} = \left[ egin{array}{cccc} m{X}_1^{
m T} & (U_1 - u_0) m{X}_1^{
m T} \ m{X}_2^{
m T} & (U_2 - u_0) m{X}_2^{
m T} \ dots & dots \ m{X}_n^{
m T} & (U_n - u_0) m{X}_n^{
m T} \end{array} 
ight], \quad m{S} = \left[ egin{array}{cccc} (m{X}_1^{
m T} & m{0}_{1 imes p}) \{m{D}_{u_1}^{
m T} m{K}_{u_1} m{D}_{u_1} \}^{-1} m{D}_{u_1}^{
m T} m{K}_{u_1} \ m{X}_2^{
m T} & m{0}_{1 imes p} \} \{m{D}_{u_2}^{
m T} m{K}_{u_2} m{D}_{u_2} \}^{-1} m{D}_{u_2}^{
m T} m{K}_{u_2} \ m{V}_2^{
m T} & m{V}_2^{
m T} m{V}_2^{$$

with  $\mathbf{0}_{1\times p}$  is the  $1\times p$  zero matrix, and

$$\mathbf{K}_{u_0} = \text{diag}\{K_h(U_1 - u_0), K_h(U_2 - u_0), \cdots, K_h(U_n - u_0)\}.$$

The solution of the problem (6) is given by

$$\left[\hat{\alpha}_{1}(u_{0}), \cdots, \hat{\alpha}_{p}(u_{0}), \hat{\alpha}'_{1}(u_{0}), \cdots, \hat{\alpha}'_{p}(u_{0})\right]^{\mathrm{T}} = \left\{D_{u_{0}}^{\mathrm{T}} K_{u_{0}} D_{u_{0}}\right\}^{-1} D_{u_{0}}^{\mathrm{T}} K_{u_{0}} (I_{n} - \rho W) Y.$$

564 GUO S, WEI C H.

Then, we have

$$\hat{\boldsymbol{\alpha}}(u_0) = (\boldsymbol{I}_p \quad \boldsymbol{0}_p) \{ \boldsymbol{D}_{u_0}^{\mathrm{T}} \boldsymbol{K}_{u_0} \boldsymbol{D}_{u_0} \}^{-1} \boldsymbol{D}_{u_0}^{\mathrm{T}} \boldsymbol{K}_{u_0} (\boldsymbol{I}_n - \rho \boldsymbol{W}) \boldsymbol{Y}$$
 (7)

where  $\mathbf{0}_{p}$  is the  $p \times p$  zero matrix.

We take  $u_0$  to be each of  $U_1, U_2, \dots, U_n$ , then we can obtain the estimators of  $\hat{\boldsymbol{\alpha}}(U_j), j = 1, 2, \dots, p$ . Then we can define the estimator for  $\boldsymbol{M}$  as

$$\bar{\boldsymbol{M}} = \left[\boldsymbol{X}_{1}^{\mathrm{T}}\hat{\boldsymbol{\alpha}}(U_{1}), \boldsymbol{X}_{2}^{\mathrm{T}}\hat{\boldsymbol{\alpha}}(U_{2}), \cdots, \boldsymbol{X}_{n}^{\mathrm{T}}\hat{\boldsymbol{\alpha}}(U_{n})\right]^{\mathrm{T}} = \boldsymbol{S}(\boldsymbol{I}_{n} - \rho \boldsymbol{W})\boldsymbol{Y}$$
(8)

Replacing M of (4) by  $\bar{M}$ , we obtain the following profile log-likelihood function

$$\log L_n(\boldsymbol{Y}|\boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) + \log|\boldsymbol{I}_n - \rho \boldsymbol{W}|$$

$$-\frac{1}{2\sigma^2} \left[ \boldsymbol{Y} - \rho \boldsymbol{W} \boldsymbol{Y} - \bar{\boldsymbol{M}} \right]^{\mathrm{T}} \left[ \boldsymbol{Y} - \rho \boldsymbol{W} \boldsymbol{Y} - \bar{\boldsymbol{M}} \right]$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) + \log|\mathbf{A}(\rho)|$$

$$-\frac{1}{2\sigma^2} \left[ (\boldsymbol{I}_n - \boldsymbol{S})(\boldsymbol{I}_n - \rho \boldsymbol{W}) \boldsymbol{Y} \right]^{\mathrm{T}} \left[ (\boldsymbol{I}_n - \boldsymbol{S})(\boldsymbol{I}_n - \rho \boldsymbol{W}) \boldsymbol{Y} \right]$$
(9)

Given  $\rho$ , by differentiating  $\log L_n(Y|\theta)$  with respect to  $\sigma^2$ , we obtain the following equation:

$$\frac{\partial \log L_n(\boldsymbol{Y}|\boldsymbol{\theta})}{\partial \sigma^2} = 0 \Rightarrow n\sigma^2 = \left[ (\boldsymbol{I}_n - \boldsymbol{S})(\boldsymbol{I}_n - \rho \boldsymbol{W}) \boldsymbol{Y} \right]^{\mathrm{T}} \left[ (\boldsymbol{I}_n - \boldsymbol{S})(\boldsymbol{I}_n - \rho \boldsymbol{W}) \boldsymbol{Y} \right]$$
 (10)

Then, the profile maximum likelihood estimator of  $\sigma^2$  can be obtained as

$$\hat{\sigma}^{2}(\rho) = \frac{1}{n} \mathbf{Y}^{\mathrm{T}} (\mathbf{I}_{n} - \rho \mathbf{W})^{\mathrm{T}} (\mathbf{I}_{n} - \mathbf{S})^{\mathrm{T}} (\mathbf{I}_{n} - \mathbf{S}) (\mathbf{I}_{n} - \rho \mathbf{W}) \mathbf{Y}$$
(11)

Furthermore, the concentrated log-likelihood function of  $\rho$  is

$$\log L_n(\rho) = -\frac{n}{2}[\log(2\pi) + 1] - \frac{n}{2}\log[\hat{\sigma}^2(\rho)] + \log|\boldsymbol{I}_n - \rho\boldsymbol{W}|$$
(12)

Maximizing  $\log L_n(\rho)$  leads to the estimator  $\hat{\rho}$  of  $\rho$ . Then, we can define the final estimator of  $\sigma^2$  as  $\hat{\sigma}^2 = \hat{\sigma}^2(\hat{\rho})$ . Substituting  $\hat{\rho}$  into (8), we can obtain the estimator of M as

$$\hat{M} = S(I - \hat{\rho}W)Y \tag{13}$$

Finally, the residual vector is

$$\hat{\varepsilon} = Y - \hat{\rho}WY - \hat{M} \tag{14}$$

## 3 Profile Generalized Likelihood Ratio Test

### 3.1 Construction of the Test Statistic

On one hand, under the alternative hypothesis  $H_1$ , we can obtain the profile log-likelihood as

$$l(H_1) = -\frac{n}{2}[\log(2\pi) + 1] - \frac{n}{2}\log[\hat{\sigma}^2(\hat{\rho})] + \log|\mathbf{I}_n - \hat{\rho}\mathbf{W}|$$
 (15)

On the other hand, if the null hypothesis  $H_0$  is true, model (1) reduces to the following standard varying coefficient model

$$Y_i = \mathbf{X}_i^{\mathrm{T}} \boldsymbol{\alpha}(U_i) + \varepsilon_i, \quad i = 1, 2, \cdots, n$$
 (16)

As a special case of model (3), we can obtain the profile log-likelihood of model (16) as

$$l(H_0) = -\frac{n}{2}[\log(2\pi) + 1] - \frac{n}{2}\log(\tilde{\sigma}^2)$$
(17)

with

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^n [Y_i - \boldsymbol{X}_i^{\mathrm{T}} \tilde{\boldsymbol{\alpha}}(U_i)]^2}{n}, \quad \tilde{\boldsymbol{\alpha}}(U_i) = (\boldsymbol{I}_p \quad \boldsymbol{0}_p) \{\boldsymbol{D}_{U_i}^{\mathrm{T}} \boldsymbol{K}_{U_i} \boldsymbol{D}_{U_i}\}^{-1} \boldsymbol{D}_{U_i}^{\mathrm{T}} \boldsymbol{K}_{U_i} \boldsymbol{Y}$$

Following Fan and Huang<sup>[13]</sup>, we define the profile generalized likelihood ratio test statistic

$$T = l(H_1) - l(H_0) = \frac{n}{2} \log \left( \frac{\tilde{\sigma}^2}{\hat{\sigma}^2(\hat{\rho})} \right) + \log |\mathbf{I}_n - \hat{\rho} \mathbf{W}|$$
 (18)

Intuitively, if  $H_0$  is true, there should not be significant difference between  $RSS(H_0)$  and  $RSS(H_1)$ . Otherwise,  $RSS(H_0) - RSS(H_1)$  will tend to take a large positive value as  $RSS(H_0)$  should become systematically larger than  $RSS(H_1)$  under  $H_1$ . Hence, a large value of the test statistic  $T_n$  indicates that the null hypothesis should be rejected. If we denote the observed value of T by t, the p-value of the test is

$$p = P_{H_0}(T \ge t) \tag{19}$$

where  $P_{H_0}(\cdot)$  refers to the probability computed under the null hypothesis  $H_0$ . For a given significance level a, if p < a then reject  $H_0$ ; otherwise not reject  $H_0$ .

#### 3.2 Calculation of the p-Value by the Residual Based Bootstrap Approach

For the proposed test statistic T, it is difficult to obtain its asymptotic null distribution as the presence of the spatial lag term in model (1). To solve this problem, the bootstrap approach was suggested to obtain the p-value by many researchers (Cai et al.<sup>[14]</sup>, Fan and Jiang<sup>[15]</sup>, Hall and Hart<sup>[16]</sup>, Herwartz and Xu<sup>[17]</sup>). In the following, we propose here a residual based bootstrap procedure to derive the p-value of the test.

**Step 1** Based on the data set  $\{Y_i, X_i, U_i\}_{i=1}^n$ , under  $H_1$ , compute the residual vector  $\hat{\boldsymbol{\varepsilon}}_1 = (\hat{\varepsilon}_{11}, \hat{\varepsilon}_{21}, \cdots, \hat{\varepsilon}_{n1})$  shown in (14) and centralize it to obtain  $\hat{\boldsymbol{\varepsilon}} = (\hat{\varepsilon}_{11} - \bar{\hat{\boldsymbol{\varepsilon}}}_1, \hat{\varepsilon}_{21} - \bar{\hat{\boldsymbol{\varepsilon}}}_1, \cdots, \hat{\varepsilon}_{n1} - \bar{\hat{\boldsymbol{\varepsilon}}}_1)$  in which  $\bar{\hat{\boldsymbol{\varepsilon}}}_1 = \frac{1}{n} \sum_{i=1}^n \hat{\varepsilon}_{i1}$ . Furthermore, with the estimation results under both  $H_0$  and  $H_1$ , compute the observed value t of the test statistic T by (18).

**Step 2** Generate the bootstrap residuals  $\hat{\boldsymbol{\varepsilon}}^* = (\hat{\varepsilon}_1^*, \hat{\varepsilon}_2^*, \cdots, \hat{\varepsilon}_n^*)$  from the empirical distribution function of  $\hat{\boldsymbol{\varepsilon}}$ . Define

$$Y_i^* = \boldsymbol{X}_i^{\mathrm{T}} \tilde{\boldsymbol{\alpha}}(U_i) + \hat{\varepsilon}_i^*$$

**Step 3** Calculate the bootstrap test statistic  $T_n^*$  based on the data set  $\{Y_i^*, X_i, U_i\}_{i=1}^n$ .

**Step 4** Repeat Steps 2 and 3 k times and obtain a bootstrap sample of the test statistic T as  $T_1^*, T_2^*, \dots, T_k^*$ . The p-value is then estimated by

$$\hat{p} = \frac{\sharp \{T_i^* : T_i^* \ge t\}}{k}$$

where t is the observed value of the test statistic T obtained in Step 1 and  $\sharp A$  denotes the number of the elements in set A.

For the sake of simplicity, we use the same bandwidth in calculating  $T_i^*$  as that in t.

566 GUO S, WEI C H.

## 4 Simulation Studies

In this section, we shall conduct some simulations to examine the performance of the proposed estimate procedure.

Assume that the observations are collected from a uniform, two-dimensional grid consisting of  $m \times m$  lattice points with unit distance between any two neighboring points along the horizontal and vertical axes. These  $m^2$  points are arranged in an orthogonal coordinate system. We generate the spatial weight matrix W according to the principle of Rook contiguity.

The data are generated from the following varying coefficient spatial autoregressive model

$$Y_i = \rho \sum_{j=1}^{n} w_{ij} Y_j + \alpha_1(u_i) x_{i1} + \alpha_2(u_i) x_{i2} + \varepsilon_i$$

where  $x_{1i} \sim U(-2,2), x_{2i} \sim N(1,1), u_i \sim U(0,1)$ , and  $\alpha_1(u_i) = \sin(2\pi u_i) + 1, \alpha_2(u_i) = 2e^{-2(2u_i-1)^2} + 3u_i$ . The spatial lag parameter  $\rho$  took each of the values in the set (-0.15, -0.1, -0.05, 0, 0.05, 0.1, 0.15). To gain an idea of the effect of the distribution of the error on our results, we take the following two different types of the error distribution, 1)  $\varepsilon_i \sim N(0, 1)$ , 2)  $\varepsilon_i \sim U(-\sqrt{3}, \sqrt{3})$ . In the simulation, the kernel function was taken to be the Gaussian kernel  $K(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2})$ , and the smoothing parameter h was taken as  $s_u n^{-1/5}$  for simplicity, where  $s_u$  is the sample standard deviation of  $u_1, u_2, \dots, u_n$ .

For each given value of  $\rho$  and each type of error distribution and bandwidth, 1000 replications with  $n=10^2, 15^2$  were run and the rejection rate at the significance level  $\alpha=0.05$  was computed as the simulated power of our proposed test procedure. And for each replication, the p-value was computed based on m=500 bootstrap samples. The results are shown in Table 1.

ρ	$\varepsilon \sim N(0,1)$		$\varepsilon \sim U(-\sqrt{3}, \sqrt{3})$	
	m = 10	m = 15	m = 10	m = 15
-0.15	0.944	0.999	0.954	1.000
-0.1	0.692	0.964	0.691	0.956
-0.05	0.229	0.453	0.245	0.464
0	0.055	0.059	0.049	0.054
0.05	0.256	0.549	0.237	0.538
0.1	0.768	0.977	0.737	0.986
0.15	0.980	1.000	0.998	1.000

**Table 1** The Rejection frequencies for  $H_0: \rho = 0$  at the significance level  $\alpha = 0.05$ 

We summarize our findings as follows. When the null hypothesis is true (that is  $\rho = 0$ ), the rejection frequencies (estimated sizes) of our proposed test are quite good and close to their nominal level 0.05 under different error distributions. Under the alternative hypothesis, the rejection rate seems very robust to the variation of the type of error distribution, and increases rapidly as the alternative hypothesis deviates from the null hypothesis.

### 5 Conclusion

In this paper, a test approach is proposed to check the existence of spatial effects in varying coefficient spatial autoregressive models, in which a residual-based bootstrap procedure is suggested to derive the *p*-value of the test. The simulation experiment demonstrates that the proposed test performs satisfactorily.

#### References

- [1] Su L J, Jin S N. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. Journal of Econometrics, 2010, 157(1): 18–33.
- [2] Li T Z, Mei C L. Testing a polynomial relationship of the non-parametric component in partially linear spatial autoregressive models. Papers in Regional Science, 2013, 92(3): 633–649.
- [3] Su L J. Semiparametric GMM estimation of spatial autoregressive models. Journal of Econometrics, 2012, 167(2): 543-560.
- [4] Cleveland W S, Grosse E, Shyu W M. Local regression models. In Statistical Models in S, Eds. by Chambers J M and Hastie T J. Wadsworth and Brooks, Pacific Grove, 1991, 309–376.
- [5] Hastie T, Tibshirani R. Varying-coefficient models (with discussion). Journal of the Royal Statistical Society Series B, 1993, 55(4): 757–796.
- [6] Fan J Q, Zhang W Y. Statistical methods with varying coefficient models. Statistics and its Interface, 2008, 1: 179–195.
- [7] Cliff A D, Ord J K. Spatial Autocorrelation. London: Pion Ltd., 1973.
- [8] Anselin L. Spatial econometrics: Methods and models. Dordrecht: Kluwer Academic Publishers, 1988.
- [9] Kelejian H H, Prucha I R. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review, 1999, 40(2): 509–533.
- [10] Lee L F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica, 2004, 72(6): 1899–1925.
- [11] Lee L F. GMM and 2SLS estimation of mixed regressive spatial autoregressive models. Journal of Econometrics, 2007, 137(2): 489–514.
- [12] Li K M, Chen J B. Profile maximum likelihood estimation of semiparametric varying coefficient spatial lag model. The Journal of Quantitative Technical Economics, 2013, 4: 85–98 (in Chinese).
- [13] Fan J Q, Huang T. Profile likelihood inferences on semiparametric varying-coefficient partially linear models. Bernoulli, 2005, 11(6): 1031–1057.
- [14] Cai Z W, Fan J Q, Yao Q W. Functional-coefficient regression models for nonlinear times series. Journal of the American Statistical Association, 2000, 95(451): 941–956.
- [15] Fan J Q, Jiang J C. Nonparametric inferences for additive models. Journal of the American Statistical Association, 2005, 100(471): 890–907.
- [16] Hall P, Hart J D. Bootstrap test for difference between means in nonparametric regression. Journal of the American Statistical Association, 1990, 85(412): 1039–1049.
- [17] Herwartz H, Xu F. A new approach to bootstrap inference in functional coefficient models. Computational Statistics and Data Analysis, 2009, 53(6): 2155–2167.