

Article

Parameter Estimation in Spatial Autoregressive Models with Missing Data and Measurement Errors

Tengjun Li , Zhikang Zhang and Yunquan Song *

College of Science, China University of Petroleum, Qingdao 266580, China; 2109050104@s.upc.edu.cn (T.L.); 2102010334@s.upc.edu.cn (Z.Z.)

* Correspondence: syqfly1980@upc.edu.cn; Tel.: +86-185-6397-1869

Abstract: This study addresses the problem of parameter estimation in spatial autoregressive models with missing data and measurement errors in covariates. Specifically, a corrected likelihood estimation approach is employed to rectify the bias in the log-maximum likelihood function induced by measurement errors. Additionally, a combination of inverse probability weighting (IPW) and mean imputation is utilized to mitigate the bias caused by missing data. Under several mild conditions, it is demonstrated that the proposed estimators are consistent and possess oracle properties. The efficacy of the proposed parameter estimation process is assessed through Monte Carlo simulation studies. Finally, the applicability of the proposed method is further substantiated using the Boston Housing Dataset.

Keywords: spatial autoregression; measurement error; corrected likelihood estimation; missing data; inverse probability weighting

MSC: 62F12; 62G08; 62G20; 62J07T07



Citation: Li, T.; Zhang, Z.; Song, Y. Parameter Estimation in Spatial Autoregressive Models with Missing Data and Measurement Errors. *Axioms* **2024**, *13*, 315. <https://doi.org/10.3390/axioms13050315>

Academic Editors: David Katz, Jiujuan Zhang and Palle E.T. Jorgensen

Received: 23 January 2024

Revised: 15 April 2024

Accepted: 24 April 2024

Published: 10 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Both classic linear regression models and spatial autoregressive models are used to study the linear relationship between a response variable and multiple explanatory variables. The former usually assumes that the observed values of the explained variable are independent of each other. However, in fields such as economics, biology, and meteorology, the collected data often exhibit certain spatial dependencies. Ignoring these dependencies in statistical inference can lead to significantly biased results (see Luo [1]). The latter model, in contrast, considers spatial dependencies, positing that a region's response variable is not only related to its explanatory variables but also associated with those of neighboring regions (see Chen [2]).

Both classic linear regression models and spatial autoregressive models typically assume that (i) the values of explanatory variables are always observable or measurable (assuming no incomplete observation), and (ii) the observations or measurements are error-free (assuming no measurement errors in explanatory variables) (see Bai et al. [3]). However, these assumptions may be violated in many scientific studies and practical applications. It is well known in statistical analysis that ignoring measurement errors and missing observations in explanatory variables can lead to serious biases in estimation and large standard errors, resulting in incorrect inference on the estimated regression coefficients. Therefore, studying parameter estimation methods for spatial autoregressive models with measurement errors and missing data in explanatory variables is of great importance.

This paper mainly studies the parameter estimation issues in spatial autoregressive models with measurement errors and missing data in explanatory variables.

Many economic datasets are related to spatial locations, such as studies on Gross Domestic Product, tourism, and research and development across various provinces nationwide (see Li [4]). Spatial data introduce spatial location information (or mutual distances)

to cross-sectional or panel data. Spatial data are generally recognized to have locational attributes (see Anselin [5]), assuming that variables with closer distances are more closely related. Tobler's First Law of Geography states that everything is related to everything else, but nearby things are more related than distant things (see Tobler [6]).

Spatial econometrics was first proposed by Jean Paelinck in May 1974 at the Netherlands Statistical Conference, aiming to provide methodological foundations for econometric models of urban and regional economics. Spatial econometrics primarily deals with addressing spatial interactions (spatial autocorrelation) and spatial structures (spatial heterogeneity) in cross-sectional and panel data regression models (see Anselin [7]). The issues studied in spatial econometrics include (1) model setting; (2) parameter estimation; (3) model setting testing; and (4) spatial prediction (see Anselin [8]). Spatial autoregressive models, which incorporate spatial effects into classic regression models, are important for studying spatial autocorrelation in data.

Research on spatial autoregressive models began early. In the 1970s, Cliff and Ord [9] introduced spatial effects into traditional linear regression models, constructing spatial autoregressive models. However, due to the endogeneity caused by spatial dependence, ordinary least-squares parameter estimates are biased and inconsistent. To address this, researchers have proposed other estimation methods to reduce or eliminate bias caused by spatial effects, obtaining consistent parameter estimates. Anselin [5] first applied the concentrated likelihood function method to provide the maximum likelihood estimation (MLE) of the model. However, MLE requires solving complex likelihood functions and is computationally expensive. Kelejian and Prucha [10] proposed the Generalized Method of Moments for spatial autoregressive models, which is relatively simpler than MLE regardless of sample size, and also provided the asymptotic properties of this estimator for large and small samples. Lesage [11] used Bayesian methods based on Gibbs sampling to address the parameter estimation issue in spatial autoregressive models with heteroskedasticity. Lee [12] employed the GMM and 2SLS for spatial autoregressive models, deriving the optimal GMM and proving its consistency and asymptotic normality. In a normal distribution, the optimal GMM and ML estimates have the same limiting distribution. The fundamental idea behind the Generalized Method of Moments (GMM) is to estimate model parameters using the moment conditions within the model. Meanwhile, Two-Stage Least Squares (2SLS) is a commonly employed method to address issues of endogeneity, particularly when instrumental variables are encountered in regression analysis. Lee and Liu [13] extended the GMM in mixed spatial autoregressive models to higher-order mixed spatial autoregressive models. Their research showed that the GMM has computational advantages over the usual ML, and the proposed GMM estimates were proven to be consistent and asymptotically normal. Wei et al. [14] proposed a semi-parametric partially linear varying coefficient spatial autoregressive model and introduced a quasi-maximum likelihood method based on local linear methods to estimate model parameters.

The development of spatial econometrics in China started late, focusing primarily on empirical studies. For example, Ma and Zhang [15] used provincial panel data in China to analyze the impact of economic development and energy structure on haze pollution, finding a positive spatial correlation in inter-provincial haze pollution, with provinces having a higher proportion of coal consumption in their energy structure experiencing more severe haze pollution. Wang et al. [16] used Chinese city panel data to analyze the impact of high-speed rail on economic growth, finding that the opening of high-speed rail strengthens the positive spatial dependence among Chinese cities' GDPs and has a positive spatial spillover effect on economic growth. Cheng and Dong [17] used panel data from countries along the "Belt and Road" to construct a spatial econometric model to analyze the spatial effects of trade facilitation on China's industrial goods exports, finding a positive spatial spillover effect. In 2010, China saw its first textbook on spatial econometrics, authored by Shen [18], but this book focused more on modeling and simulation, with less emphasis on theoretical derivation and proof. Overall, domestic research on spatial

autoregressive models has achieved certain results but generally exhibits a situation with more empirical studies and fewer theoretical methods.

Spatial data with measurement errors and missing data are commonly found in scientific research and practical applications. Although various techniques can reduce errors and missing data, the measurement errors and missing data sometimes reach a level that cannot be ignored. Therefore, studying spatial autoregressive models with measurement errors and missing data in explanatory variables becomes very important.

Due to spatial dependencies, the assumption of independence among explanatory variables in spatial data no longer holds. Under such circumstances, if one ignores the measurement errors and spatial dependencies present in the data and still employs traditional estimation methods (such as the least-squares method), this can lead to significant biases in the estimation results (see Li [4]). In their study of spatial data linear mixed models with measurement errors, Yi et al. [19] found that ignoring measurement errors leads to reduced regression coefficients and increased variance. To address this issue, they proposed a structural modeling method that obtains model parameters through maximum likelihood estimation while considering measurement errors and uses the EM algorithm for iterative optimization of parameters. They proved that the proposed method's parameter estimates have good asymptotic properties, i.e., the maximum likelihood estimates are consistent and satisfy asymptotic normality. Huque et al. [20] explored the sensitivity of parameter estimation in spatial regression models when explanatory variables have measurement errors. When errors exist, parameter estimates of the model exhibit attenuation bias. They proved the bias expression of the estimator when ignoring measurement errors, showing that the bias is related to the degree of spatial correlation between explanatory variables and residuals. They also proposed two strategies for obtaining consistent parameter estimates: (1) using an estimated attenuation factor for subsequent correction and (2) linearly transforming the error-prone explanatory variables. Through simulation studies, they assessed the finite sample performance of these two methods. The results showed that both methods can provide consistent parameter estimates, but the transformation method performs better. They also illustrated this method using ischemic heart disease data. Zhang and Zhu [21] proved that for spatial autoregressive models, when the explained variables have measurement errors, whether or not these errors are related to the model's disturbance terms, the commonly used maximum likelihood estimation is inconsistent. When the null hypothesis is rejected, using 2SLS can yield consistent parameter estimates. He and Hu [22] introduced measurement errors of independent variables into the classic spatial autoregressive model, establishing a univariate spatial autoregressive measurement error (USARME) model, and proposed a parameter estimation method for this model. Their research showed that if one does not consider the measurement errors of independent variables and directly uses the ordinary spatial autoregressive model, the estimated parameters exhibit significant bias. As measurement errors increase, the parameter estimation performance of the ordinary spatial autoregressive model becomes very poor, while the USARME model still achieves good estimation results. The feasibility and reliability of the proposed parameter estimation method were verified through numerical simulations. Luo [1] proposed a three-stage least-squares (3SLS) estimation method that simultaneously uses Berkson and classical types of instrumental variables. Under mild conditions, they derived the asymptotic normality of the estimators proposed for each type of instrumental variable.

In practical applications, for various reasons, some observations in datasets may be missing. For example, in pharmacological studies, due to the side effects of some drugs, some patients are unable to continue treatment and drop out mid-course, leading to missing data. Simply ignoring these missing values not only reduces the efficiency of the study but may also introduce systematic bias (see You [23]).

Types of missing data, classified according to the missing mechanism, can roughly be divided into three categories: Missing Completely at Random (MCAR), Missing at Random (MAR), and Missing Not at Random (MNAR). Among these, randomly missing data are characterized by the distribution of missing data not depending on unobserved

data but only on observed data (see Cheng [24]). This article mainly studies this type of missing data, that is, whether the missing data depend only on observable (exogenous) explanatory variables.

Typical methods for dealing with missing data include imputation and inverse probability weighting. Yang et al. [25] proposed a missing data imputation method based on spatial clustering and spatial autoregressive models. This method first uses the DBSCAN algorithm to cluster the dataset, then establishes a spatial autoregressive model within each cluster to impute missing data. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a density-based spatial clustering algorithm designed to organize points in a dataset into clusters separated by areas of varying density, with the capability to identify outliers or noise points. Experimental results with meteorological data showed that compared to cluster kernel function-based imputation and K-nearest neighbors (KNN) imputation, this method achieves more accurate imputation results. Wang and Lee [26] considered the situation of randomly missing explanatory variables in spatial autoregressive models. They proposed nonlinear least-squares estimation and the Generalized Method of Moments for the model. Additionally, they proposed an inferential two-stage least-squares estimation method. These estimation methods were analyzed and compared, revealing that the generalized nonlinear least-squares, best-generalized two-stage least-squares, and optimal moment estimation methods have the same asymptotic variance. Monte Carlo simulations showed that these methods provide consistent estimates even in the presence of unknown heteroskedastic disturbances and are more robust compared to the EM algorithm. Luo [1] studied the situation of missing response variable data in spatial autoregressive models. They proposed a two-stage least-squares estimation method based on IPW and missing data imputation (2II-2SLS). They proved the consistency and asymptotic normality of this estimator and studied its finite sample performance through simulations. The results showed that the performance of this estimator is superior to that of the EM estimator and the maximum likelihood estimator and that the choice of initial values has little impact on its performance.

The above-mentioned papers have greatly enhanced our understanding of parameter estimation in spatial autoregressive models with missing data and measurement errors. However, current research mainly focuses on dealing with single types of data issues in spatial autoregressive models, such as considering only missing data or only measurement errors. In contrast, there is a relative lack of research on the more complex situation where both measurement errors and missing data occur simultaneously in spatial autoregressive models. This paper investigates the parameter estimation problem in spatial autoregressive models with both measurement errors and missing data, proposing a method for parameter estimation. The main contributions of this article are as follows:

- (1) We establish a parameter estimation method for spatial autoregressive models with missing data and measurement errors, which uses a combination of corrected likelihood estimation and IPW with mean imputation to eliminate biases caused by missing data and measurement errors.
- (2) We apply the proposed method to revise and optimize traditional spatial autoregressive models. Based on this, the log-likelihood function of the modified model is presented, and explicit mathematical expressions and analyses are provided for some key parameters, offering deeper insights into the theoretical foundation and practical application of the model.
- (3) Under some mild conditions, we prove that the proposed estimates have consistency and oracle properties. Additionally, we conduct extensive numerical studies, proving that our method is superior to others in terms of parameter estimation.

The rest of this paper is organized as follows. In Section 2, the parameter estimation of spatial autoregression with missing data and measurement errors is considered, presenting the theoretical properties of the oracle estimator and proving its consistency and asymptotic normality. Section 3 conducts numerical comparisons and simulation studies. Section 4

illustrates the application of the proposed method through real data analysis. The proof of the technical results is provided in Appendix A.

2. Parameter Estimation in Spatial Autoregressive Models with Measurement Errors and Missing Data

2.1. Spatial Autoregressive Models

We consider the following spatial autoregressive model:

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

where Y_n is an $n \times 1$ vector of the observed values of the dependent variable; X_n is an $n \times k$ matrix of the observed values of k exogenous covariates; λ is a scalar spatial autoregressive coefficient with $|\lambda| < 1$; W_n is a known $n \times n$ spatial weight matrix; ε_n is an n -dimensional vector of regression disturbances, independently and identically distributed with mean 0 and finite variance σ^2 ; and β is a k -dimensional vector of the regression coefficients. Let $\theta_0 = (\sigma_0^2, \lambda_0, \beta_0^T)^T = (\theta_{1,0}, \theta_{2,0}, \dots, \theta_{k+2,0})^T$ be the true parameter point. Let $S_n(\lambda) = I_n - \lambda W_n$ and $\varepsilon_n(\delta) = Y_n - X_n\beta - \lambda W_n Y_n$, where $\delta = (\lambda, \beta^T)^T$. Then, following the approach of Lee [27], the log-likelihood function of the model is given by:

$$\begin{aligned} \ln L_n(\theta) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 + \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \varepsilon_n^T(\delta) \varepsilon_n(\delta) \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 + \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \left[(S_n(\lambda) Y_n - X_n\beta)^T (S_n(\lambda) Y_n - X_n\beta) \right] \end{aligned} \tag{2}$$

where $\theta = (\sigma^2, \lambda, \beta^T)^T = (\theta_1, \theta_2, \dots, \theta_{k+2})^T$ and $\ln L_n(\theta)$ is the log-likelihood function of model (1). Let $\varepsilon_n = (e_1, e_2, \dots, e_n)^T$, $S_n = S_n(\lambda_0)$, and $G_n = W_n S_n^{-1}$. Additionally, to ensure the large-sample properties of QMLE, some basic assumptions are listed as follows:

Assumption 1. In ε_n , e_i , $i = 1, \dots, n$ are independently and identically distributed, with mean $E(e_i) = 0$ and variance $\text{Var}(e_i) = \sigma^2$. For $\gamma > 0$, the moment $E(|e_i|^{4+\gamma})$ exists.

Assumption 2. For all i, j , the elements $w_{n,ij}$ of W_n are at most of the order of h_n^{-1} , denoted as $O(1/h_n)$, where the rate sequence h_n can be bounded or divergent. As a normalization, $w_{n,ii} = 0$ for all i .

Assumption 3. As $n \rightarrow \infty$, $n^{-1}h_n \rightarrow 0$.

Assumption 4. The matrix S_n is non-singular.

Assumption 5. The matrix sequences W_n and S_n^{-1} are uniformly bounded in terms of row and column sums.

Assumption 6. For all n , elements X_n are uniformly bounded constants. When $\lim_{n \rightarrow \infty} n^{-1} X_n^T X_n$ exists and is non-singular.

Assumption 7. For all λ in the compact parameter space Λ , which is a compact set of the parameter space Λ , $S_n^{-1}(\lambda)$ is uniformly bounded in terms of row and column sums. The true λ_0 is inside Λ .

Assumption 8. $\lim_{n \rightarrow \infty} n^{-1} (X_n, G_n X_n \beta_0)^T (X_n, G_n X_n \beta_0)$ exists and is a non-singular matrix.

Assumption 9. $\lim_{n \rightarrow \infty} E \left(n^{-1} \frac{\partial^2 \ln L_n(\theta_0)}{\partial \theta \partial \theta^T} \right)$ exists.

Assumption 10. For all θ in the open set H containing the true parameter point θ_0 , the third-order derivative $\frac{\partial^3 \ln L_n(\theta)}{\partial \theta_j \partial \theta_l \partial \theta_m}$ exists. Additionally, there exists a function M_{jlm} such that for all $\theta \in H$, $\left| n^{-1} \frac{\partial^3 \ln L_n(\theta)}{\partial \theta_j \partial \theta_l \partial \theta_m} \right| \leq M_{jlm}$, where $E(M_{jlm}) < \infty$ for all j, l, m .

Assumptions 1–9 are similar to those provided by Lee [27], which are sufficient conditions for the correctness of the global identification and the consistency and asymptotic normality of QMLE for model (1). Assumption 1 is needed to apply the central limit theorem by Kelejian and Prucha [28]. Assumptions 2 and 3 characterize the weight matrix for sample size n . If h_n is a bounded sequence, then Assumptions 2 and 3 are satisfied. In Case’s model, Assumptions 2 and 3 still hold, although h_n may diverge (see Case [29]). Assumption 4 is used to ensure the existence of the means and variances of the independent variables. Assumption 5 implies that when n tends to infinity, the variance of Y_n is bounded (refer to Kelejian, Prucha [28], and Lee [27]). Assumption 6 excludes multicollinearity among the regressors in X_n . For convenience, we assume the regressors are uniformly bounded. If not, they can be replaced with random regressors under certain finite moment conditions (see Lee [27]). Assumption 7 is meaningful for handling the nonlinearity of the log-likelihood function $\ln|S_n(\lambda)|$. Assumptions 8 and 9 are applicable to the asymptotic normality of QMLE. Assumption 10 is similar to condition (C) in Fan and Li [30] and plays an important role in the Taylor expansion of related functions.

2.2. Spatial Autoregressive Model with Missing Data and Measurement Errors

When a subset of covariates has missing values, we consider model (3). Let $X_i^{(o)} \in \mathbb{R}^s$ be the vector of covariates that are always observed and $X_i^{(m)} \in \mathbb{R}^k$ be the vector of covariates that may contain some missing components. For each observation, the indicator Q_i denotes whether $X_i^{(m)}$ is fully observed, i.e., if $X_i^{(m)}$ is fully observed, then $Q_i = 1$; otherwise, $Q_i = 0$. Let $v_i = (X_i^{(o)})^T \in \mathbb{R}^s$, and as mentioned earlier, the data in this study are randomly missing but not endogenous. This means that the probability of missing observations may depend on variables that are always observed rather than on variables that may have missing data. Formally:

$$\pi_{i_0} = \Pr(Q_i = 1 \mid X_i) = \Pr(Q_i = 1 \mid X_i^{(o)}) = \Pr(Q_i = 1 \mid v_i) \tag{3}$$

In the covariate data, we assume that the dimension of the covariates $X_i^{(m)} \in \mathbb{R}^k$ with missing data and the dimension of the covariates involved in the missing model $X_i^{(o)} \in \mathbb{R}^s$ are fixed.

For the issue of missing data in covariates, the IPW method can be used to address this. The idea of IPW is to offset potential biases due to missing data by assigning different weights to complete observations. This approach helps mitigate biases in results estimation due to missing covariate data, thereby achieving more accurate statistical inferences. The probability π_{i_0} is usually parameterized and modeled through logistic or probit regression. Here, we assume it is generated by the following logistic regression model:

$$\tilde{\pi}_i = \frac{\exp(\xi_0 + v_i^T \xi_1)}{1 + \exp(\xi_0 + v_i^T \xi_1)} \tag{4}$$

For simplicity, π_{i_0} denotes the true probability of observation i having complete data, and $\tilde{\pi}_i$ is the probability calculated based on the logistic function. Let $Q = \text{diag}\left(\frac{Q_1}{\tilde{\pi}_1}, \frac{Q_2}{\tilde{\pi}_2}, \dots, \frac{Q_n}{\tilde{\pi}_n}\right)$. The weighted spatial autoregressive log-likelihood function is defined as:

$$\ln \tilde{L}_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 + \ln|S_n(\lambda)| - \frac{1}{2\sigma^2} \left[(S_n(\lambda)Y_n - X_n\beta)^T Q (S_n(\lambda)Y_n - X_n\beta) \right] \tag{5}$$

When X_n has measurement errors, consider the classic additive measurement error model:

$$Z_n = X_n + U_n \tag{6}$$

In Equation (6), X_n cannot be directly observed, but Z_n can be, where Z_n is an $n \times k$ matrix, U_n is the error term with $U_n \sim N(0, \Sigma)$, and ε_n is independent of U_n . Additionally, we assume that U_n is independent of the covariate X_n .

Li [4] proposed a corrected likelihood estimation to solve the spatial autoregressive model with measurement errors. In this case, IPW and corrected likelihood estimation are applied to spatial autoregressive models with measurement errors and missing data.

The corrected likelihood method was initially proposed by Nakamura [31] to address the impact of measurement errors on parameter estimation, enabling parameter estimation without additional assumptions. The specific method is as follows:

For the classic linear regression model with measurement errors:

$$\begin{cases} Y = X\beta + \varepsilon, \\ Z = X + U \end{cases} \tag{7}$$

where Y is the dependent variable, X is the unobservable value of the explanatory variable, β is the parameter vector, ε is the residual vector, Z is the observable vector, and U is the measurement error.

Let $L(\beta, X, Y)$, $U(\beta, X, Y)$, $J(\beta, X, Y)$, and $I(\beta, X, Y)$ denote, respectively, the log-likelihood function, score function, observed information, and Fisher information of model (7) given Z and Y , with E^+ being the mathematical expectation regarding the respective variable Y . Without measurement errors in variables, the following equations hold:

$$E^+U(\beta, X, Y) = 0 \tag{8}$$

$$E^+J(\beta, X, Y) = I(\beta, X, Y) \tag{9}$$

With measurement errors, if Z values are simply substituted for X , Equations (8) and (9) do not always hold.

Thus, Nakamura’s corrected likelihood method is used to handle this model, setting the corrected log-likelihood function $L^*(\beta, Z, Y)$ to satisfy:

$$E^*\{L^*(\beta, Z, Y)\} = L(\beta, X, Y) \tag{10}$$

where E^* is the conditional expectation of Z given Y, X . Let $U^*(\beta, Z, Y) = \frac{\partial L^*(\beta, Z, Y)}{\partial \beta}$ and $J^*(\beta, Z, Y) = -\frac{\partial U^*(\beta, Z, Y)}{\partial \beta}$ represent the corrected score function and corrected observed information, respectively. If E^* and $\partial\beta$ are interchangeable, then:

$$E^*\{U^*(\beta, Z, Y)\} = U(\beta, X, Y) \tag{11}$$

$$E^*\{J^*(\beta, Z, Y)\} = I(\beta, X, Y) \tag{12}$$

If the estimation of β satisfies $U^*(\beta^*, Z, Y) = 0$, then β^* is called a corrected likelihood estimate. Let $E = E^+E^*$. Then,

$$E\{U^*(\beta, Z, Y)\} = E^+E^*\{U^*(\beta, Z, Y)\} = E^+\{U(\beta, X, Y)\} = 0 \tag{13}$$

This shows that the corrected score function is unbiased.

Property 1. Let F be an open convex subset of the parameter space containing β . If $L^*(\beta, Z, Y)$ and $L(\beta, X, Y)$ are differentiable on F , $\sum k^{-2}var\{L^*(\beta, Z_k, Y_k)\} < \infty$, β is identifiable, Y is mutually identifiable, and then $U^*(\beta, X, Y) = 0$ has a root that is consistent with probability one as $n \rightarrow \infty$ converges in probability to 0.

Property 2. If β^* and β_X are consistent roots of $U^*(\beta, X, Y) = 0$ and $U(\beta, Z, Y) = 0$, and if $U^*(\beta, X, Y)$ and $U(\beta, Z, Y)$ meet some regularity conditions, then under the given conditions of X and Y , $(\beta^* - \beta_X)$ follows an asymptotic normal distribution with mean 0 and variance $I^+(\beta, X)^{-1}E^+[var\{U^*(\beta, Z, Y)\}]I^+(\beta, X)^{-1}$, as $(n \rightarrow \infty)$.

For model (7), the log-likelihood function is

$$L(\beta, X, Y) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} (Y_n - X_n\beta)^T (Y_n - X_n\beta) \tag{14}$$

Define

$$L^*(\beta, Z, Y) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} [(Y_n - X_n\beta)^T (Y_n - X_n\beta) - \beta^T \Sigma \beta] \tag{15}$$

Then,

$$E^*\{L^*(\beta, Z, Y)\} = L(\beta, X, Y) \tag{16}$$

Thus, $L^*(\beta, Z, Y)$ is the corrected likelihood function for model (7).

By applying Nakamura’s corrected likelihood method, the corrected likelihood function for the spatial autoregressive model with missing data and measurement errors can be obtained (here, denote $\ln \hat{L}_n(\theta) = L^*(\lambda, \beta, Z, Y)$):

$$\ln \hat{L}_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 + \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \{ [(S_n(\lambda)Y_n - Z_n\beta)^T Q(S_n(\lambda)Y_n - Z_n\beta)] - \beta^T \Sigma \beta \} \tag{17}$$

Next, we solve for the parameters. Differentiating both sides of the previous equation with respect to β and σ^2 , we obtain the following set of equations:

$$\begin{cases} \frac{\partial \ln \hat{L}_n(\theta)}{\partial \beta} = 0 \\ \frac{\partial \ln \hat{L}_n(\theta)}{\partial \sigma^2} = 0 \end{cases} \tag{18}$$

That is,

$$\begin{cases} -\frac{1}{2\sigma^2} [-2Z_n^T Q(S_n(\lambda)Y_n - Z_n\beta) - 2\Sigma\beta] = 0 \\ -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} [(S_n(\lambda)Y_n - X_n\beta)^T Q(S_n(\lambda)Y_n - X_n\beta) - \beta^T \Sigma \beta] = 0 \end{cases} \tag{19}$$

Given λ , the corrected likelihood estimate for β is:

$$\hat{\beta}(\lambda) = (Z_n^T Q Z_n - \Sigma)^{-1} Z_n^T Q S_n(\lambda) Y_n \tag{20}$$

Similarly, the corrected likelihood estimate for σ^2 is:

$$\hat{\sigma}^2(\lambda) = \frac{1}{n} \{ [(S_n(\lambda)Y_n - Z_n\beta)^T Q(S_n(\lambda)Y_n - Z_n\beta)] - \beta^T \Sigma \beta \} = \frac{1}{n} Y_n^T S_n^T(\lambda) M_n S_n(\lambda) Y_n \tag{21}$$

where $M_n = P_z - QZ_n \left((Z_n^T Q Z_n - \Sigma)^{-1} \right)^T \Sigma (Z_n^T Q Z_n - \Sigma)^{-1} Z_n^T Q$, $P_z = Q - QZ_n (Z_n^T Q Z_n - \Sigma)^{-1} Z_n^T Q$. Substituting Equations (20) and (21) into Equation (17), the concentrated likelihood function for λ is:

$$\ln \hat{L}_n(\lambda) = -\frac{n}{2} (\ln(2\pi) + 1) - \frac{n}{2} \ln \hat{\sigma}^2(\lambda) + \ln |S_n(\lambda)| \tag{22}$$

The corrected likelihood estimate of λ is then found by maximizing Equation (22).

Theorem 1 (Oracle Properties). *Suppose the regularity conditions in Assumptions A1–A5 in Appendix A hold. It is apparent that Assumptions A1–A5 are largely consistent with the earlier Assumptions 1–8. If Assumptions A1–A5 hold, then $\hat{\theta}_n$ is globally identified, a consistent estimator, and has an asymptotic distribution.*

1. θ_0 is globally identifiable, and $\hat{\theta}_n$ is a consistent estimate of θ_0 .
2. $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{D} N(0, \Sigma_\theta^{-1})$, where $\Sigma_\theta^{-1} = -\lim_{n \rightarrow \infty} E\left(\frac{1}{n} \frac{\partial^2 \ln \hat{L}_n(\theta)}{\partial \theta \partial \theta^T}\right)$.

3. Simulation

Through Monte Carlo simulation, the performance and efficiency of the proposed method were compared. We simulated 500 datasets.

The sample sizes n for each dataset were set to 100, 150, 200, and 250, respectively. The threshold m of the weight matrix represents the number of non-zero elements in each row of the matrix. The threshold m for the weight matrix was set to 10, 15, 20, and 25, respectively. The spatial autoregressive coefficients λ were set to 0.5 and -0.5 . Covariates and random errors were generated as follows: Covariates $X_n = (X_1, X_2, \dots, X_p)$ were generated from a p -dimensional normal distribution with mean 0 and variance 1. In the simulation, we set $\beta = (1, 0, 5, 0, 3)$. The generation mechanism for Y_n is

$$Y_n = (I_n - \lambda W_n)^{-1}(X_n \beta + \varepsilon_i) \tag{23}$$

We assumed that the error ε in the spatial autoregressive model followed a normal distribution with variances e^2 of $0.5^2, 1^2$, and 1.5^2 . In the simulation, we assumed that X_2 and X_4 might have missing values. The missingness model considered is

$$\text{Logit}\{Pr(R_i = 1)\} = 0.5 + 0.5X_{i_1} - 1.5X_{i_3} + 0.5X_{i_5} \tag{24}$$

The measurement error model was set as $Z = X + U$, where the measurement error U follows a multivariate normal distribution Σ with mean 0 and variance 1^2 .

- (I) Both measurement errors and missing data are considered;
- (II) Measurement errors are ignored;
- (III) Missing data are ignored (By dropping observations with missing covariates.);
- (IV) Both measurement errors and missing data are ignored.

Tables 1–3 present the median square errors (MeSE) of the estimates for λ , β , and σ^2 as $(\hat{\lambda} - \lambda)^2$, $(\hat{\beta} - \beta)^2$, and $(\hat{\sigma}^2 - \sigma^2)^2$, respectively. As observed from Tables 1–3, the estimates provided by the proposed method are overall significantly superior to those obtained by directly ignoring missing data or both missing data and measurement errors, with smaller squared errors. Compared to ignoring missing data, the proposed method’s correction effect on estimating β and λ is not very pronounced, but the correction for σ^2 is relatively significant. Overall, the proposed method more significantly corrects the biases caused by missing data. However, when the values of n and e are relatively low, the correction effect for measurement errors is not particularly satisfactory. Moreover, ignoring both missing data and measurement errors leads to larger squared errors due to the severe loss of information. In summary, correcting for missing data and measurement errors is both necessary and effective.

Table 1. MeSEs for λ .

e	λ	$n = 100$			$n = 150$			$n = 200$			$n = 250$		
		$m = 10$	$m = 15$	$m = 20$	$m = 10$	$m = 15$	$m = 20$	$m = 15$	$m = 20$	$m = 25$	$m = 15$	$m = 20$	$m = 25$
Incorporating measurement errors and missing data													
0.5	0.5	1.627E+01	9.324E-01	5.839E+01	4.618E-02	4.189E-02	4.618E-02	3.850E-02	3.492E-02	3.850E-02	2.462E-02	3.809E-02	2.610E-02
	-0.5	8.240E+00	2.289E+00	3.887E+00	4.440E-02	4.288E-02	4.440E-02	3.938E-02	2.922E-02	3.938E-02	2.435E-02	3.501E-02	2.692E-02
1	0.5	2.184E-01	2.354E-01	2.102E-01	1.552E-01	1.353E-01	1.555E-01	1.487E-01	1.248E-01	1.185E-01	1.078E-01	1.304E-01	1.073E-01
	-0.5	2.081E-01	2.065E-01	2.159E-01	1.480E-01	1.437E-01	1.562E-01	1.524E-01	1.312E-01	1.207E-01	1.138E-01	1.281E-01	1.112E-01
1.5	0.5	4.731E-01	4.785E-01	4.959E-01	3.910E-01	3.290E-01	3.192E-01	3.803E-01	2.695E-01	2.539E-01	2.395E-01	2.561E-01	2.375E-01
	-0.5	4.632E-01	5.174E-01	4.847E-01	3.844E-01	3.290E-01	3.267E-01	3.723E-01	2.694E-01	2.819E-01	2.261E-01	2.668E-01	2.542E-01
Ignoring missing data													
0.5	0.5	1.320E+02	1.603E+02	1.368E+03	9.288E+02	2.599E+02	8.833E+01	8.905E+02	8.535E+01	9.025E+01	7.097E+01	6.560E+01	1.963E+02
	-0.5	1.403E+02	1.288E+02	6.390E+02	1.014E+03	2.685E+02	8.620E+01	9.785E+02	1.797E+02	1.074E+02	7.247E+01	6.491E+01	1.774E+02
1	0.5	3.652E+03	6.698E+01	1.661E+03	7.315E+02	2.568E+02	9.291E+01	6.350E+02	1.839E+02	1.098E+02	7.515E+01	7.648E+01	1.835E+02
	-0.5	3.840E+03	7.060E+01	1.649E+03	1.035E+03	3.026E+02	8.620E+01	5.711E+02	1.797E+02	1.081E+02	7.477E+01	7.326E+01	1.934E+02
1.5	0.5	3.776E+03	6.636E+01	2.115E+03	9.394E+02	2.887E+02	9.215E+01	7.766E+02	1.727E+02	1.091E+02	7.751E+01	7.263E+01	1.949E+02
	-0.5	3.692E+03	6.373E+01	1.862E+03	8.851E+02	2.952E+02	8.631E+01	7.435E+02	1.987E+02	1.167E+02	7.406E+01	6.734E+01	1.884E+02
Ignoring measurement errors													
0.5	0.5	5.054E-02	5.442E-02	5.727E-02	3.850E-02	3.523E-02	3.850E-02	3.140E-02	3.492E-02	2.867E-02	2.482E-02	3.501E-02	2.432E-02
	-0.5	5.584E-02	4.883E-02	4.941E-02	3.938E-02	3.660E-02	3.938E-02	2.713E-02	2.922E-02	2.754E-02	2.403E-02	3.501E-02	2.692E-02
1	0.5	2.101E-01	2.141E-01	2.100E-01	1.334E-01	1.487E-01	1.332E-01	1.199E-01	1.248E-01	1.185E-01	1.074E-01	1.314E-01	1.095E-01
	-0.5	2.021E-01	2.038E-01	2.149E-01	1.428E-01	1.524E-01	1.490E-01	1.064E-01	1.312E-01	1.207E-01	1.109E-01	1.254E-01	1.102E-01
1.5	0.5	4.449E-01	4.604E-01	4.668E-01	3.803E-01	3.180E-01	3.142E-01	2.530E-01	2.695E-01	2.539E-01	2.441E-01	2.509E-01	2.375E-01
	-0.5	4.664E-01	4.983E-01	4.969E-01	3.723E-01	3.233E-01	3.121E-01	2.527E-01	2.694E-01	2.819E-01	2.309E-01	2.637E-01	2.568E-01
Ignoring measurement errors and missing data													
0.5	0.5	1.227E+02	1.543E+02	1.312E+03	8.905E+02	2.516E+02	8.535E+01	6.129E+02	1.699E+02	9.582E+01	6.998E+01	6.491E+01	1.921E+02
	-0.5	1.299E+02	1.220E+02	6.024E+02	9.785E+02	2.573E+02	8.620E+01	6.237E+02	1.942E+02	1.035E+02	7.140E+01	6.491E+01	1.746E+02
1	0.5	3.468E+03	6.335E+01	1.557E+03	7.074E+02	2.468E+02	8.876E+01	6.189E+02	1.827E+02	1.078E+02	7.358E+01	7.555E+01	1.799E+02
	-0.5	3.493E+03	6.724E+01	1.557E+03	1.008E+03	2.891E+02	8.326E+01	5.566E+02	1.772E+02	1.062E+02	7.408E+01	7.166E+01	1.914E+02
1.5	0.5	3.524E+03	6.336E+01	2.019E+03	8.924E+02	2.822E+02	8.807E+01	7.539E+02	1.677E+02	1.078E+02	7.674E+01	7.112E+01	1.918E+02
	-0.5	3.375E+03	6.215E+01	1.734E+03	8.151E+02	2.867E+02	8.371E+01	7.322E+02	1.935E+02	1.143E+02	7.346E+01	6.638E+01	1.847E+02

Table 2. MeSEs for β .

e	λ	$n = 100$			$n = 150$			$n = 200$			$n = 250$		
		$m = 10$	$m = 15$	$m = 20$	$m = 10$	$m = 15$	$m = 20$	$m = 15$	$m = 20$	$m = 25$	$m = 15$	$m = 20$	$m = 25$
Incorporating measurement errors and missing data													
0.5	0.5	1.801E-02	8.973E-04	6.942E-03	1.063E-07	2.664E-07	1.269E-06	5.714E-08	2.531E-07	6.536E-07	4.296E-07	2.436E-06	2.780E-07
	−0.5	2.281E-02	2.045E-03	1.019E-03	1.119E-07	3.246E-07	1.563E-06	5.643E-08	1.796E-07	5.257E-07	4.646E-07	2.419E-06	2.300E-07
1	0.5	1.784E-07	1.216E-05	5.359E-07	2.187E-07	9.488E-07	3.077E-06	1.683E-07	7.737E-07	2.334E-06	1.877E-06	3.835E-06	9.066E-07
	−0.5	1.822E-07	1.164E-05	4.690E-07	2.171E-07	9.527E-07	3.141E-06	1.917E-07	8.920E-07	1.713E-06	1.825E-06	4.568E-06	1.024E-06
1.5	0.5	4.202E-07	2.680E-05	1.040E-06	3.780E-07	1.599E-06	3.857E-06	4.070E-07	1.672E-06	4.307E-06	4.402E-06	5.968E-06	2.078E-06
	−0.5	4.165E-07	2.693E-05	8.912E-07	3.953E-07	1.725E-06	6.557E-06	4.159E-07	1.475E-06	4.581E-06	4.567E-07	5.814E-06	2.216E-06
Ignoring missing data													
0.5	0.5	1.027E-06	1.165E-06	1.946E-07	6.254E-08	1.896E-07	9.387E-07	5.386E-08	2.548E-07	6.431E-07	4.499E-07	2.257E-06	2.740E-07
	−0.5	1.329E-06	1.328E-06	2.800E-07	6.503E-08	2.004E-07	9.951E-07	5.913E-08	1.693E-07	5.375E-07	4.490E-07	2.253E-06	2.388E-07
1	0.5	1.679E-07	1.181E-05	5.124E-07	2.171E-07	9.079E-07	3.141E-06	1.828E-07	7.743E-07	2.231E-06	2.041E-06	3.937E-06	9.629E-07
	−0.5	1.866E-07	1.037E-05	4.678E-07	2.425E-07	9.527E-07	3.857E-06	1.939E-07	8.920E-07	1.685E-06	1.822E-06	3.700E-06	1.017E-06
1.5	0.5	4.402E-07	2.404E-05	1.753E-06	3.953E-07	1.725E-06	6.790E-06	4.070E-07	1.654E-06	3.690E-06	4.567E-07	5.917E-06	1.985E-06
	−0.5	4.567E-07	2.572E-05	9.034E-07	4.730E-07	1.725E-06	6.706E-06	4.159E-07	1.475E-06	4.581E-06	4.402E-07	5.814E-06	2.232E-06
Ignoring measurement errors													
0.5	0.5	1.801E-02	8.973E-04	6.942E-03	1.063E-07	2.664E-07	1.269E-06	5.714E-08	2.531E-07	6.536E-07	4.296E-07	2.436E-06	2.780E-07
	−0.5	2.281E-02	2.045E-03	1.019E-03	1.119E-07	3.246E-07	1.563E-06	5.643E-08	1.796E-07	5.257E-07	4.646E-07	2.419E-06	2.300E-07
1	0.5	1.784E-07	1.216E-05	5.359E-07	2.187E-07	9.488E-07	3.077E-06	1.683E-07	7.737E-07	2.334E-06	1.877E-06	3.835E-06	9.066E-07
	−0.5	1.822E-07	1.164E-05	4.690E-07	2.171E-07	9.527E-07	3.141E-06	1.917E-07	8.920E-07	1.713E-06	1.825E-06	4.568E-06	1.024E-06
1.5	0.5	4.202E-07	2.680E-05	1.040E-06	3.780E-07	1.599E-06	3.857E-06	4.070E-07	1.672E-06	4.307E-06	4.402E-06	5.968E-06	2.078E-06
	−0.5	4.165E-07	2.693E-05	8.912E-07	3.953E-07	1.725E-06	6.557E-06	4.159E-07	1.475E-06	4.581E-06	4.567E-07	5.814E-06	2.216E-06
Ignoring measurement errors and missing data													
0.5	0.5	1.801E-02	8.973E-04	6.942E-03	1.063E-07	2.664E-07	1.269E-06	5.714E-08	2.531E-07	6.536E-07	4.296E-07	2.436E-06	2.780E-07
	−0.5	2.281E-02	2.045E-03	1.019E-03	1.119E-07	3.246E-07	1.563E-06	5.643E-08	1.796E-07	5.257E-07	4.646E-07	2.419E-06	2.300E-07
1	0.5	1.784E-07	1.216E-05	5.359E-07	2.187E-07	9.488E-07	3.077E-06	1.683E-07	7.737E-07	2.334E-06	1.877E-06	3.835E-06	9.066E-07
	−0.5	1.822E-07	1.164E-05	4.690E-07	2.171E-07	9.527E-07	3.141E-06	1.917E-07	8.920E-07	1.713E-06	1.825E-06	4.568E-06	1.024E-06
1.5	0.5	4.202E-07	2.680E-05	1.040E-06	3.780E-07	1.599E-06	3.857E-06	4.070E-07	1.672E-06	4.307E-06	4.402E-06	5.968E-06	2.078E-06
	−0.5	4.165E-07	2.693E-05	8.912E-07	3.953E-07	1.725E-06	6.557E-06	4.159E-07	1.475E-06	4.581E-06	4.567E-07	5.814E-06	2.216E-06

Table 3. MeSEs for σ^2 .

e	λ	$n = 100$			$n = 150$			$n = 200$			$n = 250$		
		$m = 10$	$m = 15$	$m = 20$	$m = 10$	$m = 15$	$m = 20$	$m = 15$	$m = 20$	$m = 25$	$m = 15$	$m = 20$	$m = 25$
Incorporating measurement errors and missing data													
0.5	0.5	8.838E+03	6.371E+00	1.498E+05	2.544E-02	2.725E-02	2.938E-02	1.402E-02	1.326E-02	1.413E-02	1.255E-02	5.831E-02	1.436E-02
	−0.5	1.971E+03	3.881E+01	4.972E+02	2.515E-02	2.724E-02	2.161E-02	1.504E-02	1.285E-02	1.285E-02	1.087E-02	4.399E-02	1.340E-02
1	0.5	2.497E-01	2.111E-01	2.422E-01	2.286E-01	2.451E-01	2.327E-01	2.655E-01	3.116E-01	2.646E-01	5.731E-01	7.831E-01	5.634E-01
	−0.5	2.428E-01	2.424E-01	2.191E-01	2.618E-01	2.378E-01	2.618E-01	2.571E-01	3.228E-01	2.817E-01	5.984E-01	1.015E+00	5.455E-01
1.5	0.5	1.039E+00	8.811E-01	8.929E-01	1.870E+00	1.237E+00	1.388E+00	1.751E+00	1.950E+00	1.752E+00	3.281E+00	4.377E+00	2.882E+00
	−0.5	7.428E-01	1.214E+00	1.008E+00	2.276E+00	1.156E+00	1.197E+00	1.409E+00	2.525E+00	2.000E+00	3.183E+00	5.228E+00	3.731E+00
Ignoring missing data													
0.5	0.5	6.195E+05	7.860E+05	9.643E+07	9.921E+07	1.059E+07	6.249E+05	1.097E+08	6.918E+06	1.673E+06	1.590E+06	1.887E+06	9.845E+06
	−0.5	7.591E+05	6.743E+05	1.409E+07	1.534E+08	1.149E+07	8.153E+05	8.139E+06	1.796E+06	8.070E+06	1.241E+06	1.750E+06	8.070E+06
1	0.5	5.165E+08	8.849E+04	1.396E+08	8.536E+07	1.048E+07	8.269E+05	1.222E+08	7.143E+06	1.647E+06	1.489E+06	1.885E+06	8.864E+06
	−0.5	8.202E+08	8.025E+04	1.175E+08	1.300E+08	9.951E+06	7.452E+05	1.033E+08	7.292E+06	1.689E+06	1.381E+06	1.962E+06	1.062E+07
1.5	0.5	5.896E+08	9.098E+04	1.807E+08	1.364E+08	1.323E+07	7.982E+05	1.827E+08	8.376E+06	1.992E+06	1.701E+06	2.093E+06	9.215E+06
	−0.5	5.829E+08	8.085E+04	1.362E+08	1.450E+08	1.148E+07	7.943E+05	1.349E+08	8.539E+06	1.874E+06	1.478E+06	2.036E+06	1.101E+07
Ignoring measurement errors													
0.5	0.5	1.552E-02	1.409E-02	1.635E-02	2.452E-02	1.870E-02	1.839E-02	3.074E-02	3.737E-02	3.452E-02	3.967E-02	1.438E-01	5.327E-02
	−0.5	1.556E-02	1.611E-02	1.420E-02	1.966E-02	1.673E-02	1.157E-01	3.432E-02	3.214E-02	5.016E-02	4.095E-02	1.157E-01	5.016E-02
1	0.5	1.575E-01	1.784E-01	1.942E-01	3.701E-01	4.468E-01	3.903E-01	4.643E-01	5.033E-01	4.763E-01	7.663E-01	1.166E+00	7.863E-01
	−0.5	2.203E-01	1.996E-01	2.201E-01	4.989E-01	3.800E-01	3.196E-01	4.460E-01	5.350E-01	5.115E-01	8.282E-01	1.174E+00	7.716E-01
1.5	0.5	1.071E+00	9.059E-01	1.040E+00	2.490E+00	1.723E+00	1.843E+00	2.196E+00	2.450E+00	2.661E+00	3.956E+00	4.853E+00	3.374E+00
	−0.5	8.711E-01	1.130E+00	1.075E+00	2.974E+00	1.652E+00	1.563E+00	1.849E+00	3.028E+00	2.583E+00	3.761E+00	5.851E+00	4.455E+00
Ignoring measurement errors and missing data													
0.5	0.5	6.249E+05	7.933E+05	9.505E+07	1.007E+08	1.088E+07	6.285E+05	1.143E+08	8.101E+06	1.670E+06	1.592E+06	1.891E+06	1.008E+07
	−0.5	8.039E+05	6.779E+05	1.376E+07	1.524E+08	1.148E+07	8.325E+05	8.101E+06	1.797E+06	8.078E+06	1.262E+06	1.754E+06	8.078E+06
1	0.5	5.243E+08	8.745E+04	1.387E+08	8.553E+07	1.051E+07	8.309E+05	1.223E+08	7.152E+06	1.655E+06	1.490E+06	1.887E+06	8.866E+06
	−0.5	8.338E+08	8.011E+04	1.187E+08	1.306E+08	1.015E+07	7.443E+05	1.036E+08	7.452E+06	1.691E+06	1.382E+06	1.965E+06	1.069E+07
1.5	0.5	5.856E+08	8.961E+04	1.826E+08	1.371E+08	1.389E+07	7.982E+05	1.828E+08	8.344E+06	1.997E+06	1.726E+06	2.096E+06	9.126E+06
	−0.5	5.829E+08	8.209E+04	1.369E+08	1.456E+08	1.156E+07	7.943E+05	1.357E+08	8.559E+06	1.843E+06	1.474E+06	2.039E+06	1.089E+07

4. Real Data Example

In this section, we present a real example to illustrate the performance of the parameter estimation procedure for spatial autoregressive models with missing data and measurement errors proposed in this paper.

We consider the famous 1970 Boston Housing Dataset, which contains information on 506 different houses in different locations in the Boston Standard Metropolitan Statistical Area. This dataset has been used by many authors and can be found in the `spdep` library in R. It was first analyzed by Harrison and Rubinfeld [32]. Sun et al. [33] and Du et al. [34] explored the spatial dependence of these data through partially linear varying coefficient autoregressive and partially linear additive autoregressive models, respectively. Liu [35] used the Moran I statistic to test the spatial dependence of the dataset. Therefore, the data serve our analysis purposes well. Table 4 provides specific descriptions of the variables in the dataset.

Table 4. Descriptions of the variables in the Boston Housing Dataset.

Attribute	Explanation	Remarks
CRIM	Per capita crime rate by town	
ZN	Proportion of residential land zoned for lots over 25,000 sq. ft.	Residential land proportion
INDUS	Proportion of non-retail business acres per town	Non-retail business proportion
CHAS	Charles River dummy variable	Charles River variable for regression analysis
NOX	Nitrogen oxide concentration (ppm)	Environmental indicator
RM	Average number of rooms per dwelling	Number of rooms in residential units
AGE	Proportion of owner-occupied units built prior to 1940	Pre-1940s-constructed units proportion
DIS	Weighted distances to five Boston employment centers	Distance to employment hubs
RAD	Index of accessibility to radial highways	Highway accessibility index
TAX	Full-value property tax rate per 10,000	Property tax rate
PRATO	Pupil–teacher ratio by town	Pupil-teacher ratio
B	$1000(Bk - 0.63)^2$, where Bk is the proportion of blacks by town	Proportion of black population
LSTAT	Percentage of the population classified as lower income	Lower-income class proportion
MEDV	Median value of owner-occupied homes	Typically, the target variable in an analysis

In the actual data analysis, following the practice of Harrison and Rubinfeld [32], we consider the logarithm of the median value of owner-occupied homes (MEDV) in census tracts as the dependent variable and the other variables as the independent variables. Among these, the weighted distance to five Boston employment centers (DIS), the index of accessibility to radial highways (RAD), the percentage of the population classified as lower income (LSTAT), and the average number of rooms per dwelling (RM) are log-transformed, while the nitrogen oxide concentration (NOX) is squared. For ease of analysis, all variables are mean-centered to have a sample mean of zero.

Spatial weight matrices generally consist of two types of information. One is determined using latitude and longitude coordinates, and the other is determined using the relative locations of regions (see Liu [35]).

Our approach is similar to that of Pace and Gilley [36]. We first define an initial matrix W and the weight between two houses i and j as:

$$W_{ij} = \max\left(1 - \frac{d_{ij}}{d_0}, 0\right) \tag{25}$$

where d_{ij} is the Euclidean distance calculated based on the latitude and longitude coordinates of the two houses. We set the threshold distance d_0 to 0.05. Additionally, in practice, the spatial weight matrix is row-normalized.

For the above dataset, we consider the following model:

$$Y = \lambda WY + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + X_5\beta_5 + \varepsilon_i \tag{26}$$

where the response variable Y is the median value of house prices (MEDV), X_1 is the weighted distance (DIS), X_2 is the index of accessibility to radial highways (RAD), X_3 is the percentage of the population classified as lower income (LSTAT), X_4 is the nitrogen oxide concentration (NOX), and X_5 is the average number of rooms per dwelling (RM).

In the simulation, we assume that X_2 and X_4 may have missing values. The missingness model considered is:

$$\text{Logit}\{Pr(R_i) = 1\} = 1 + 2X_{i1} - 3X_{i3} + X_{i5} \tag{27}$$

The measurement error model is set as $Z = X + U$, where the measurement error U follows a multivariate normal distribution Σ with mean 0 and variance equal to the sample variance σ_i^2 of each variable.

Table 5 displays the median squared errors (MeSEs) of the estimators λ , β , and σ^2 , specifically $||\hat{\lambda} - \lambda||^2$, $||\hat{\beta}_i - \beta_i||^2$, and $||\hat{\sigma}^2 - \sigma^2||^2$. It can be observed that the performance of the proposed method for the Boston Housing Dataset is largely consistent with the results of the numerical simulations. Although the correction effect for the partial β is not particularly pronounced, the overall correction effect is still relatively ideal, and in particular, the corrections for σ^2 and λ are very effective.

Table 5. MeSEs for λ , β , and σ^2 .

	λ	β_1	β_2	β_3	β_4	β_5	σ^2
Real parameters	4.446E-01	-1.770E-01	-3.425E-02	-4.261E-01	-1.249E-01	4.102E-01	8.168E-03
Incorporating measurement errors and missing data	8.870E-03	1.878E-03	8.669E-04	9.953E-03	1.716E-03	2.848E-01	9.182E-07
Ignoring missing data	3.536E-02	2.045E-03	8.998E-04	9.691E-03	1.805E-03	1.746E-01	8.075E-06
Ignoring measurement errors	1.043E-02	1.919E-03	8.840E-04	6.065E-03	2.132E-03	2.457E-01	8.077E-07
Ignoring measurement errors and missing data	4.528E-02	2.094E-03	8.565E-04	3.985E-03	1.629E-03	1.208E-01	1.125E-05

5. Conclusions

We have developed a robust method for simultaneously handling missing data and measurement errors in covariates of spatial autoregressive models. Clearly, traditional statistical methods may lead to biased estimates when covariates have missing data and measurement errors. Our method uses IPW and corrected likelihood methods to address this issue. We have studied the theoretical properties of the proposed method and investigated its performance in parameter estimation through Monte Carlo simulations, comparing it to scenarios where measurement errors, missing data, or both are ignored. The simulation studies demonstrate that our method outperforms traditional direct extensions for parameter estimation in spatial autoregressive models.

Author Contributions: Software, Validation, Investigation, Resources, Data Curation, Writing—Review & Editing, and Funding Acquisition were handled by T.L. He was mainly responsible for implementing the concepts, programming and running simulations, testing the code, collecting and curating data, as well as editing the manuscript and securing funding. Conceptualization, Methodology, Formal Analysis, Writing—Original Draft Preparation, Writing—Review & Editing were performed by Z.Z. He was primarily responsible for the initial concept and design, theoretical derivation of formulas, and writing and editing the theoretical parts of the manuscript. Supervision and Project Administration were carried out by the corresponding author Y.S., who was in charge of overseeing the entire project and managing its administration. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Fundamental Research Funds for the Central Universities (No. 23CX03012A) and the National Key Research and Development Program (2021YFA1000102) of China.

Data Availability Statement: All datasets used in this study, including those analyzed and those generated during the research process, are publicly available and can be accessed through the following link: <https://drive.google.com/drive/folders/1egSNneFuDZe-iUbn69OPkWIDAxEzP6q> (accessed on 25 April 2024).

Conflicts of Interest: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Appendix A

Using a combination of IPW and mean interpolation, consider an ideal “pseudo-complete dataset” \tilde{Z}_n for covariates, satisfying

$$(S_n(\lambda)Y_n - Z_n\beta)^T Q(S_n(\lambda)Y_n - Z_n\beta) = (S_n(\lambda)Y_n - \tilde{Z}_n\beta)^T (S_n(\lambda)Y_n - \tilde{Z}_n\beta) \tag{A1}$$

From Equations (1), (4) and (A1), we obtain

$$Y_n = \lambda W_n Y_n + \tilde{Z}_n \beta + \varepsilon_n - U_n \beta \triangleq \lambda W_n Y_n + \tilde{Z}_n \beta + \varepsilon_n^* \tag{A2}$$

where $\varepsilon_n^* = \varepsilon_n - U_n \beta$.

Thus, Equation (16) can be simplified as:

$$Y_n = S_n^{-1}(\tilde{Z}_n \beta_0 + \varepsilon_n^*) \tag{A3}$$

where $S_n = S_n(\lambda_0)$.

Let $G_n = W_n S_n^{-1}$. Then, $I_n + \lambda_0 G_n = S_n^{-1}$, and Equation (A1) can be expressed as:

$$Y_n = \lambda_0 G_n \tilde{Z}_n \beta_0 + \tilde{Z}_n \beta_0 + S_n^{-1} \varepsilon_n^* \tag{A4}$$

To establish the asymptotic properties of the estimator, the following regularity conditions are required:

Assumption A1. In $\varepsilon_n = \{e_i\}$ and $U_n = \{u_i\}$, elements $e_i, u_i, i = 1, \dots, n$ are independently and identically distributed with mean $E(e_i) = E(u_i) = 0$, variance $\text{Var}(e_i) = \sigma^2, \text{Var}(u_i) = \Sigma$. For $\gamma > 0$, moments $E(\{|e_i|^{4+\gamma}\})$ and $E(\{|u_i|^{4+\gamma}\})$ exist.

Assumption A2. For all i, j , elements $w_{n,ij}$ of W_n are at most of the order of h_n^{-1} , denoted as $O(1/h_n)$, where the rate sequence $\{h_n\}$ can be bounded or divergent. For normalization, all i have $w_{n,ii} = 0$.

Assumption A3. The matrix sequences $\{W_n\}$ and $\{S_n^{-1}\}$ are uniformly bounded in terms of row and column sums.

Assumption A4. For all n , elements \tilde{Z}_n are uniformly bounded constants. When $\lim_{n \rightarrow \infty}, n^{-1} \tilde{Z}_n^T \tilde{Z}_n$ exists and is non-singular.

Assumption A5. For all λ in the compact parameter space $\Lambda, \{S_n^{-1}(\lambda)\}$ are uniformly bounded in terms of row and column sums. The true λ_0 is inside Λ .

Assumption A6. $\lim_{n \rightarrow \infty} n^{-1} (\tilde{Z}_n G_n \tilde{Z}_n \beta_0)^T (\tilde{Z}_n G_n \tilde{Z}_n \beta_0)$ exists and is a non-singular matrix.

Theorem A1. Under Assumptions 1–8, θ_0 is globally identifiable, and $\hat{\theta}_0$ is a consistent estimator of θ_0 .

Define $P_n(\lambda) = \max_{\beta, \sigma^2} E(\ln \hat{L}_n(\theta))$. In this maximization problem, the optimal solutions for β and σ^2 are

$$\begin{aligned} \beta^*(\lambda) &= \left(\tilde{Z}_n^T \tilde{Z}_n - \Sigma \right)^{-1} \tilde{Z}_n^T S_n(\lambda) Y_n \\ \sigma^{*2}(\lambda) &= \frac{1}{n} E \left[(S_n(\lambda) Y_n - \tilde{Z}_n \beta)^T (S_n(\lambda) Y_n - \tilde{Z}_n \beta) - \beta^T \Sigma \beta \right] \\ &= \frac{1}{n} \left[(\lambda_0 - \lambda)^2 (G_n \tilde{Z}_n \beta_0)^T M_n (G_n \tilde{Z}_n \beta_0) + \sigma_0^2 \text{tr} \left[(S_n^T)^{-1} S_n^T(\lambda) S_n(\lambda) S_n^{-1} \right] \right] \end{aligned}$$

Then,

$$P_n(\lambda) = -\frac{n}{2} (\ln(2\pi) + 1) - \frac{n}{2} \ln \hat{\sigma}^{*2}(\lambda) + \ln |S_n(\lambda)|$$

The value of λ_0 can be obtained by maximizing $\left\{ \frac{P_n(\lambda)}{n} \right\}$.

To prove that $\hat{\theta}_0$ is a consistent estimator of θ_0 , it suffices to show the following:

1. $\frac{\ln \hat{L}(\theta) - P_n(\lambda)}{n}$ converges uniformly to 0 in Λ ;
 2. For $\forall \omega > 0$, $\lim_{n \rightarrow \infty} \max_{\lambda \in \bar{N}_\omega(\lambda_0)} \frac{P_n(\lambda) - P_n(\lambda_0)}{n}$ is a complement set of a neighborhood with a diameter.
- (a) Since

$$\frac{\ln \hat{L}(\lambda) - P_n(\lambda)}{n} = -\frac{1}{2} \left(\ln \hat{\sigma}^2(\lambda) - \ln \sigma^{*2}(\lambda) \right)$$

$\sigma^{*2}(\lambda)$ and $\hat{\sigma}^2(\lambda)$ can be written as

$$\begin{aligned} \sigma^{*2}(\lambda) &= \frac{(\lambda_0 - \lambda)^2 (G_n \tilde{Z}_n \beta_0)^T M_n (G_n \tilde{Z}_n \beta_0)}{n} + \sigma^2(\lambda), \\ \hat{\sigma}^2(\lambda) &= \frac{1}{n} Y_n^T S_n^T(\lambda) M_n S_n(\lambda) Y_n \\ &= \frac{1}{n} (\lambda_0 - \lambda)^2 (G_n \tilde{Z}_n \beta_0)^T M_n (G_n \tilde{Z}_n \beta_0) + 2(\lambda_0 - \lambda) H_{1n}(\lambda) + H_{2n}(\lambda). \end{aligned}$$

where

$$\begin{aligned} \sigma^2(\lambda) &= \frac{\sigma_0^2}{n} \text{tr} \left[(S_n^T)^{-1} S_n^T(\lambda) S_n(\lambda) S_n^{-1} \right], \\ H_{1n}(\lambda) &= \frac{1}{n} (G_n \tilde{Z}_n \beta_0)^T M_n S_n(\lambda) S_n^{-1} \epsilon_n^*, \\ H_{2n}(\lambda) &= \frac{1}{n} \epsilon_n^{*T} (S_n^{-1})^T S_n^T(\lambda) M_n S_n(\lambda) S_n^{-1} \epsilon_n^*. \end{aligned}$$

It can be shown that on Λ ,

$$\begin{aligned} H_{1n}(\lambda) &= O_p(1), \\ H_{2n}(\lambda) - \sigma^2(\lambda) &= O_p(1). \end{aligned}$$

Therefore,

$$\hat{\sigma}^2(\lambda) - \sigma^{*2}(\lambda) = O_p(1).$$

Hence,

$$\sup_{\lambda \in \Lambda} \left| \frac{\ln \hat{L}_n(\theta) - P_n(\lambda)}{n} \right| = O_p(1).$$

(b)

$$\frac{1}{n}[P_n(\lambda) - P_n(\lambda_0)] = \frac{1}{n}[P_{p,n}(\lambda) - P_{p,n}(\lambda_0)] = \frac{1}{2}[\ln \sigma^{*2}(\lambda) - \ln \sigma^2(\lambda)],$$

where $P_{p,n}(\lambda) = -\frac{n}{2}(\ln(2\pi) + 1) - \frac{n}{2} \ln \sigma^2(\lambda) + \ln |S_n(\lambda)|.$

$\{\frac{P_n(\lambda)}{n}\}$ is uniformly continuous on Λ . By Jensen’s inequality, for all λ ,

$$\frac{P_{p,n}(\lambda) - P_{p,n}(\lambda_0)}{n} \leq 0$$

and, therefore,

$$\sigma^{*2}(\lambda) \geq \sigma^2(\lambda)$$

If the assumption does not hold, then there exists a sequence $\{\lambda_n\} \in \Lambda, \lim_{n \rightarrow \infty} \lambda_n = \lambda_+ \neq \lambda_0$ such that

$$\lim_{n \rightarrow \infty} \frac{P_n(\lambda_n) - P_n(\lambda_0)}{n} = 0$$

This can only occur if

$$\lim_{n \rightarrow \infty} (\sigma^{*2}(\lambda_+) - \sigma^2(\lambda_+)) = 0$$

and

$$\lim_{n \rightarrow \infty} \frac{P_{p,n}(\lambda_+) - P_{p,n}(\lambda_0)}{n} = 0$$

simultaneously hold, but the latter equation contradicts

$$\lim_{n \rightarrow \infty} \frac{(G_n \tilde{Z}_n \beta_0)^T M_n (G_n \tilde{Z}_n \beta_0)}{n} \neq 0$$

Hence, consistency is proven.

Theorem A2. The asymptotic distribution of CMLE $\hat{\theta}_n$ can be derived from the Taylor expansion of $\frac{\partial \ln \hat{L}_n(\hat{\theta}_n)}{\partial \theta} = 0$ at θ_n , where the first-order derivative of the log-likelihood function at θ_0 is

$$\begin{aligned} \frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \beta} &= \frac{1}{\sqrt{n}\sigma_0^2} (\tilde{Z}_n^T Q \varepsilon_n^* + n \Sigma \beta_0) \\ \frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \sigma^2} &= \frac{1}{2\sqrt{n}\sigma_0^4} (Q \varepsilon_n^{*T} \varepsilon_n^* - n \beta_0^T \Sigma \beta_0 - n \sigma_0^2) \\ \frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \lambda} &= \frac{1}{\sqrt{n}\sigma_0^2} (G_n \tilde{Z}_n \beta_0)^T \varepsilon_n^* + \frac{1}{\sqrt{n}\sigma_0^2} (\varepsilon_n^{*T} G_n \varepsilon_n^* - \sigma_0^2 \text{tr}(G_n)) \end{aligned}$$

The asymptotic distribution of these expressions can be obtained through the central limit theorem. If $\{h_n\}$ is a bounded sequence, Kelejian and Prucha’s central limit theorem can be used. If $\{h_n\}$ is unbounded, i.e., $\lim_{n \rightarrow \infty} h_n = \infty$, then under Assumption A5, $\frac{1}{\sqrt{n}\sigma_0^2} (G_n \tilde{Z}_n \beta_0)^T \varepsilon_n^*$ will significantly influence $\frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \lambda}$. This is because

$$\begin{aligned} \text{var} \left(\frac{1}{\sqrt{n}} \varepsilon_n^{*T} G_n \varepsilon_n^* \right) &= O \left(\frac{1}{h_n} \right) \\ \frac{1}{\sqrt{n}} (\varepsilon_n^{*T} G_n \varepsilon_n^* - \sigma_0^2 \text{tr}(G_n)) &= O_p(1) \end{aligned}$$

However,

$$\frac{1}{\sqrt{n}}(G_n \tilde{Z}_n \beta_0)^T \varepsilon_n^* = O_p(1)$$

In this case, Kolmogorov’s central limit theorem can be used.

$$\frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \theta}$$

’s variance matrix is $E\left(\frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \theta} \frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \theta}\right) = -E\left(\frac{1}{n} \frac{\partial^2 \ln \hat{L}_n(\theta_0)}{\partial \theta \partial \theta^T}\right)$

where

$$-E\left(\frac{1}{n} \frac{\partial^2 \ln \hat{L}_n(\theta_0)}{\partial \theta \partial \theta^T}\right) = \begin{pmatrix} \frac{1}{n\sigma_0^2}(\tilde{Z}_n^T \tilde{Z}_n - 2n\Sigma) & \frac{1}{n\sigma_0^2} \tilde{Z}_n^T (G_n \tilde{Z}_n \beta_0) & \frac{1}{\sigma_0^4} \Sigma \beta_0 \\ \frac{1}{n\sigma_0^2} (G_n \tilde{Z}_n \beta_0)^T \tilde{Z}_n & \frac{1}{n\sigma_0^2} (G_n \tilde{Z}_n \beta_0)^T (G_n \tilde{Z}_n \beta_0) + \frac{1}{n} \text{tr}(G_n^T G_n) & \frac{1}{n\sigma_0^2} (G_n) \\ \frac{1}{\sigma_0^4} \Sigma \beta_0 & \frac{1}{n\sigma_0^2} \text{tr}(G_n) & \frac{1}{2\sigma_0^4} \end{pmatrix}$$

$\ln \frac{\partial^2 \ln \hat{L}_n(\theta_0)}{\partial \theta \partial \theta^T}$, λ , β , and $\frac{1}{\sigma_0^2}$ appear either as linear or quadratic moments, and

$$\frac{\partial^2 \ln \hat{L}(\theta)}{\partial \lambda^2} = -\text{tr}\left([W_n^T S_n^{-1}(\lambda)]^2\right) - \frac{Y_n^T W_n^T W_n Y_n}{\sigma^2}$$

Let $G_n(\lambda) = W_n S_n^{-1}(\lambda)$, and by the mean value theorem,

$$\text{tr}(G_n^2(\tilde{\lambda}_n)) = \text{tr}(G_n^2) + 2\text{tr}(G_n^3(\tilde{\lambda}))(\tilde{\lambda} - \lambda_0)$$

Assumption A3 ensures that within a neighborhood of λ_0 , $G_n(\tilde{\lambda}_n)$ ’s row and column sums are uniformly bounded. Since $\text{tr}(G_n^3(\tilde{\lambda}_n)) = O\left(\frac{n}{h_n}\right)$,

$$Y_n^T W_n^T W_n Y_n = O_p(nh_n)$$

and, therefore,

$$\frac{1}{n} \left[\frac{\partial^2 \ln \hat{L}_n(\tilde{\theta}_n)}{\partial \lambda^2} - \frac{\partial^2 \ln \hat{L}_n(\theta_0)}{\partial \lambda^2} \right] = -2 \frac{\text{tr}(G_n^3(\tilde{\lambda}_n))}{n} (\tilde{\lambda}_n - \lambda_0) + \left[\frac{1}{\sigma_0^2} - \frac{1}{\tilde{\sigma}^2} \right] \frac{Y_n^T W_n^T W_n Y_n}{n} = O_p(1)$$

The second-order derivatives of other terms can be similarly deduced. Since both $\frac{\tilde{Z}_n^T G_n \varepsilon_n^*}{n}$ and $\frac{1}{n}(\varepsilon_n^{*T} G_n \varepsilon_n^* - (\sigma_0^2 + \beta_0^T \Sigma \beta_0) \text{tr}(G_n))$ converge to $O_p(1)$, it follows that

$$\frac{1}{n} \left[\frac{\partial^2 \ln \hat{L}_n(\tilde{\theta}_n)}{\partial \theta \partial \theta^T} - \frac{\partial^2 \ln \hat{L}_n(\theta_0)}{\partial \theta \partial \theta^T} \right] \xrightarrow{P} 0$$

Since $\frac{\partial \ln L_n(\theta_0)}{\partial \theta}$ is either a linear or quadratic function of ε_n^* , and the higher-order moments of ε_n^* exist according to Assumption A1, the central limit theorem by Kelejian and Prucha implies

$$\frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \theta} \xrightarrow{D} N(0, \Sigma_\theta)$$

Assumption A5 ensures that Σ_θ is non-singular, and the asymptotic distribution of $\hat{\theta}_n$ can be written as

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = - \left(\frac{1}{n} \frac{\partial^2 \ln \hat{L}_n(\tilde{\theta}_n)}{\partial \theta \partial \theta^T} \right)^{-1} \frac{1}{\sqrt{n}} \frac{\partial \ln \hat{L}_n(\theta_0)}{\partial \theta}$$

where $\hat{\theta}_n$ converges in probability to θ_0 .

References

1. Luo, G. *Statistical Inference in Spatial Autoregressive Models with Complex Data*; Beijing University of Technology: Beijing, China, 2023. [[CrossRef](#)]
2. Chen, Q. *Advanced Econometrics and Stata Application*; Higher Education Press: Beijing, China, 2010.
3. Bai, Y.; Tian, M.; Tang, M.-L.; Lee, W.-Y. Variable selection for ultra-high dimensional quantile regression with missing data and measurement error. *Stat. Methods Med. Res.* **2021**, *30*, 129–150. [[CrossRef](#)] [[PubMed](#)]
4. Li, W. *Parameter Estimation of Spatial Autoregressive Models with Measurement Error*; Yunnan University: Kunming, China, 2020.
5. Anselin, L. *SpaceStat Tutorial: A Workbook for Using SpaceStat in the Analysis of Spatial Data*; West Virginia University: Urbana, IL, USA, 1992.
6. Tobler, R.W. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* **2016**, *46*, 234–240. [[CrossRef](#)]
7. Anselin, L. Spatial Econometrics: Methods and Models. *J. Am. Stat. Assoc.* **1990**, *85*, 160. [[CrossRef](#)]
8. Anselin, L. Thirty Years of Spatial Econometrics. *Pap. Reg. Sci.* **2010**, *89*, 3–25. [[CrossRef](#)]
9. Cox, T.F. Spatial Processes: Models and Applications. *J. R. Stat. Soc. Ser. A* **1984**, *147*, 515. [[CrossRef](#)]
10. Prucha, K.I.R. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *Int. Econ. Rev.* **2010**, *40*, 509–533. [[CrossRef](#)]
11. Lesage, J.P. Bayesian Estimation of Spatial Autoregressive Models. *Int. Reg. Sci. Rev.* **1997**, *20*, 113–129. [[CrossRef](#)]
12. Lee, L.F. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *J. Econ.* **2007**, *137*, 489–514. [[CrossRef](#)]
13. Lee, L.F.; Liu, X. Efficient GMM estimation of high order spatial autoregressive models. *Econ. Theory* **2010**, *26*, 187–230. [[CrossRef](#)]
14. Wei, C.; Guo, S.; Zhai, S. Statistical inference of partially linear varying coefficient spatial autoregressive models. *Econ. Model.* **2017**, *64*, 553–559. [[CrossRef](#)]
15. Ma, M.; Zhang, X. Spatial Effects of China's Haze Pollution and the Impact of Economy and Energy Structure. *China Ind. Econ.* **2014**, 19–31. [[CrossRef](#)]
16. Wang, Y.F.; Ni, P.F. Economic Growth Spillovers and Regional Spatial Optimization Under the Influence of High-Speed Rail. *China Ind. Econ.* **2016**, 21–36. [[CrossRef](#)]
17. Cheng, Y.; Dong, C. Study on the Spatial Effects of Trade Facilitation on China's Industrial Manufactured Goods Export. *Quant. Econ. Tech. Econ. Res.* **2021**, *38*, 98–115. [[CrossRef](#)]
18. Shen, T.; Feng, D.; Sun, T. *Spatial Econometrics*; Peking University Press: Beijing, China, 2010.
19. Yi, L.; Tang, H.; Lin, X. Spatial Linear Mixed Models with Covariate Measurement Errors. *Stat. Sin.* **2009**, *19*, 1077.
20. Huque, M.H.; Bondell, H.D.; Ryan, L. On the impact of covariate measurement error on spatial regression modelling. *Environmetrics* **2015**, *25*, 560–570. [[CrossRef](#)]
21. Zhang, Z.; Zhu, P. Estimation and Testing of Spatial Autoregressive Models with Measurement Error. *Stat. Res.* **2010**, *27*, 103–109. [[CrossRef](#)]
22. He, X.; Hu, X. Parameter Estimation of Univariate Spatial Autoregressive Measurement Error Models. *Sci. China Math.* **2020**, *50*, 613–628. [[CrossRef](#)]
23. You, P. *Estimation of Semiparametric Spatial Autoregressive Models with Random Missing*; Yunnan University: Kunming, China, 2018.
24. Cheng, D. *Modeling and Analysis of Residents' Travel Behavior Based on Multiple Differences*; Dalian Jiaotong University: Dalian, China, 2019. [[CrossRef](#)]
25. Yang, Z.; Yu, J.; Chen, J. A Missing Data Imputation Method Based on Clustering and Spatial Autoregressive Model. Intelligent Information Technology Application Association. In Proceedings of the 2011 International Conference on Ecological Protection of Lakes-Wetlands-Watershed and Application of 3S Technology (EPLWW3S 2011 V2), Nanchang, China, 25–26 June 2011; pp. 554–557.
26. Wang, W.; Lee, L. Estimation of spatial autoregressive models with randomly missing data in the dependent variable. *Econ. J.* **2013**, *16*, 73–102. [[CrossRef](#)]
27. Lee, L.F. Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica* **2004**, *72*, 1899–1925. [[CrossRef](#)]
28. Kelejian, H.H.; Prucha, I.R. On the Asymptotic Distribution of the Moran I Test Statistic with Applications. *J. Econom.* **2001**. [[CrossRef](#)]
29. Case, A.C. Spatial Patterns in Household Demand. *Econometrica* **1991**, *59*, 953–965. [[CrossRef](#)]
30. Li, F.R. Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Publ. Am. Stat. Assoc.* **2001**, *96*, 1348–1360. [[CrossRef](#)]
31. Nakamura, T. Corrected Score Function for Errors-in-Variables Models: Methodology and Application to Generalized Linear Models. *Biometrika* **1990**, *77*, 127–137. [[CrossRef](#)]
32. David, H., Jr. Hedonic housing prices and the demand for clean air. *J. Environ. Econ. Manag.* **1978**. [[CrossRef](#)]
33. Sun, Y.; Yan, H.; Zhang, W. A semiparametric spatial dynamic model. *Ann. Stats.* **2014**, *42*, 700–727. [[CrossRef](#)]
34. Du, J.; Sun, X.; Cao, R. Statistical inference for partially linear additive spatial autoregressive models. *Spat. Stat.* **2018**, 52–67. [[CrossRef](#)]

35. Liu, X.; Chen, J.; Cheng, S. A Penalized Quasi-Maximum Likelihood Method for Variable Selection in the Spatial Autoregressive Model. *Spat. Stat.* **2018**, *25*, 86–104. [[CrossRef](#)]
36. Pace, R.K.; Gilley, O.W. Using the Spatial Configuration of the Data to Improve Estimation. *J. Real Estate Financ. Econ.* **1997**, *14*, 333–340. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.