Inference in Spatial Econometric Models

Niklas Ahlgren* Linda Gerkman†

January 18, 2007

Abstract

In spatial econometrics the problem of stationarity has not received much attention. Typically, the specification of models has been based on the geographical arrangement of observations, or on contiguity. In this paper we give conditions for when a spatial econometric model is stationary, based on the theory of unilateral spatial autoregressive processes. The stationarity condition cannot be checked in the model, since the model is a reduced form model in which the parameter space is restricted to the stationary region. We study the properties of the estimator of the autoregressive parameter. In the stationary case the estimator is consistent and asymptotically normal. In the non-stationary case the estimator is inconsistent and diverges. This explains the findings in some empirical papers with values of the spatial autoregressive coefficient close to one. The finite sample properties of the estimator are investigated by simulation. Finally, we study model specification and the impact of the choice of weights matrix on the estimator. An application to house prices is provided.

Key words: SPATIAL AUTOCORRELATION COEFFICIENT, SPATIAL AUTOREGRESSIVE MODEL, SPATIAL UNILATERAL PROCESS, WEIGHTS MATRIX.

*Swedish School of Economics, Department of Finance and Statistics (`niklas.ahlgren@shh.fi`).

†Swedish School of Economics, Department of Finance and Statistics (`linda.gerkman@shh.fi`).

1 Introduction

Assume that there exists spatial dependence between observations on a variable y_{ij}. Then some of the variation in y_{ij} is then explained by neighboring observations. In spatial econometrics, data is of the form $\mathbf{Y} = [y_{ij}]$, where $i = 1, \ldots, N_1$, $j = 1, \ldots, N_2$ and $\mathbf{Y}$ is an $N_1 \times N_2$ matrix of observations. The data can e.g. be generated by a spatial autoregressive (SAR) process (see Section 2).

The most commonly used model is the spatial autoregressive (SAR) model

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{y} = [y_{ij}]$ is an $N \times 1$ vector of observations on a variable y_{ij}, $\mathbf{W} = [w_{ij}]$ is an $N \times N$ weights matrix, ρ is a scalar parameter and $\boldsymbol{\varepsilon} = [\varepsilon_{ij}]$ is an $N \times 1$ vector of error terms. The simplest assumption about $\boldsymbol{\varepsilon}$ is that its elements are independent and identically distributed (IID) with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. The model can be solved in terms of $\boldsymbol{\varepsilon}$ as follows

$$\mathbf{y} = (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}, \tag{2}$$

provided that $(\mathbf{I} - \rho \mathbf{W})$ is non-singular, from which it is seen that the errors of the reduced form model are not IID. The model can be estimated by maximum likelihood (ML) by maximising the log likelihood function. Then under some regularity conditions the ML estimator is consistent and symptomatically normal.

In this paper we consider inference in the spatial autoregressive model (1) when data are generated by a unilateral spatial autoregressive (SAR) process. In spatial econometrics, as opposed to spatial statistics and the probability literature, the problem of stationarity has not received much attention. The specification of models has been based on the assumption that the elements of the weights matrix are nonstochastic and exogenous to the model. Typically, they are based on the geographical arrangement of the observations, or on contiguity (Anselin 2003, p. 313). By restricting attention to a narrower class of unilateral SAR processes, it is possible to study the problems of stationarity and inference in the framework of the SAR model. From the probability literature we give conditions for when the SAR model is stationary, based on the theory of unilateral spatial autoregressive processes (see e.g. Basu and Reinsel 1993 and Yao and Brockwell 2006). The stationarity condition cannot be checked in the model, since the model is a reduced form model in which the parameter space is restricted to the stationary region. We study the properties of the estimator of the autoregressive parameter. In the stationary case the estimator is consistent and asymptotically normal. In the non-stationary case the estimator is inconsistent and diverges. This explains the findings in some empirical papers with values of the spatial autoregressive coefficient close to one. The finite sample properties of the estimator are investigated by simulation. Finally, we study model specification and the impact of the choice of weights matrix on the estimator. An application to house prices is provided.

The paper is structured as follows. Section 2 reviews the theory of spatial unilateral autoregressive processes. Section 3 introduces the spatial econometric model. Section 4 investigates the properties of the estimator of the autoregressive parameter. Section 5 deals with extensions of the model. Section 6 contains an empirical application to house prices in the county of Stockholm, Sweden. Section 7 concludes.

We use the following notation. The matrix $\mathbf{Y} = [y_{ij}]$ is an $N_1 \times N_2$ matrix of observations, $\mathbf{y} = \text{vec}(\mathbf{Y}')$ denotes the $N \times 1$ row vectorisation of the matrix $\mathbf{Y}$ and $N = N_1 N_2$. The matrix $\mathbf{W}_i$ denotes a general weights matrix and the subindex i refers to the number of parameters in the SAR process or the number of neighbours.

2 SPATIAL UNILATERAL AR PROCESSES

The spatial unilateral autoregressive (SAR) process in two dimensions is given by (Basu and Reinsel 1993)

$$y_{ij} = \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} y_{i-k,j-l} + \varepsilon_{ij}, \quad \alpha_{00} = 0, \tag{3}$$

$\{\varepsilon_{ij}\} \sim \text{IID}(0, \sigma^2)$. In this model, introduced by Whittle (1954) and studied in detail by Tjøstheim (1978, 1981, 1983), the value at the site (i, j) is a finite autoregression on the values at the sites which lie in the lower quarter plane, i.e. y_{ij} is an autoregression on y_{il}, $l < j$, and y_{kl}, $k < i$, l unrestricted.

Tjøstheim (1978, Theorem 5.1) gives the stationarity condition for the unilateral SAR process in n dimensions. For $n = 2$ and writing $\mathbf{z} = (z_1, z_2)$, define the characteristic polynomial

$$\Phi(z_1, z_2) = 1 - \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} z_1^k z_2^l. \quad (4)$$

The process is stationary if

$$\Phi(z_1, z_2) \neq 0 \quad \text{for all } |z_1| \leq 1, |z_2| \leq 1, \quad (5)$$

i.e., the roots of $\Phi(z_1, z_2) = 0$ all lie outside the two-dimensional polydisc. For the unilateral spatial AR process we therefore have an easily checked condition for stationarity.

We consider two special cases of the general model (3). Following Basu and Reinsel (1993), if in (3) $p_1 = p_2$ (i.e. the only non-zero coefficients are α_{10}, α_{01} and α_{11}), then we obtain the special case of the first-order unilateral SAR model

$$y_{ij} = \alpha_1 y_{i-1,j} + \alpha_2 y_{i,j-1} + \alpha_3 y_{i-1,j-1} + \varepsilon_{ij}. \quad (6)$$

This model is studied in detail by Basu and Reinsel (1993). For the first-order spatial unilateral process (6), with $\Phi(z_1, z_2) = 1 - \alpha_1 z - \alpha_2 z_2 - \alpha_3 z_1 z_2$, the stationarity condition is that none of the roots of $\Phi(z_1, z_2) = 0$ lie within the closed polydisc ($|z_1| \leq, |z_2| \leq 1$). Basu and Reinsel (1993) give in Proposition 1 a set of conditions on the SAR parameters α_1, α_2 and α_3 under which the process is stationary. The conditions are:

(i) $|\alpha_i| < 1$ for all i,
(ii) $(1 + \alpha_1^2 - \alpha_2^2 - \alpha_3^2)^2 - 4(\alpha_1 + \alpha_2\alpha_3)^2 > 0$,
(iii) $1 - \alpha_2^2 > |\alpha_1 + \alpha_2\alpha_3|$.

If $\alpha_3 = 0$ in (6), i.e., the model is

$$y_{ij} = \alpha_1 y_{i-1,j} + \alpha_2 y_{i,j-1} + \varepsilon_{ij}, \quad (7)$$

then the conditions (i)–(iii) reduce to $|\alpha_1| + |\alpha_2| < 1$. This condition was derived by Whittle (1954) and Besag (1972). If $|\alpha_1| + |\alpha_2| = 1$, the model contains a unit root. See Baran et al. (2005) for unit roots in SAR processes.

We close this section by remarking that inference in spatial autoregressive (and moving average, ARMA) processes can be based on the Gaussian likelihood function (see Yao and Brockwell 2006 and the references cited therein).

In spatial econometrics, inference is based on somewhat different and simpler models, introduced already in Section 1, and to which we turn in the next section. Nevertheless, the theory of spatial unilateral AR processes is useful for understanding the properties of these models as well.

3 SPATIAL ECONOMETRIC MODELS

The spatial autoregressive (SAR) model can be written as

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \boldsymbol{\varepsilon}. \tag{8}$$

Here $\mathbf{y} = [y_{ij}]$ is an $N \times 1$ vector of observations y_{ij}, $\mathbf{W} = [w_{ij}]$ is an $N \times N$ weights matrix, ρ is a scalar parameter and $\boldsymbol{\varepsilon} = [\varepsilon_{ij}]$ is an $N \times 1$ vector of errors, which are assumed to be independent and identically distributed with mean 0 and variance σ^2. The covariance matrix of $\boldsymbol{\varepsilon}$ is of the form $\sigma^2 \mathbf{I}_N$.

The weights matrix $\mathbf{W}$ is assumed to be exogenous to the model. However, every weights matrix $\mathbf{W}$ implies an underlying probability structure for the observations $\mathbf{y}$. For simplicity, we assume that $N_1 = N_2 = n$ and $N = N_1 N_2 = n^2$. In general, the weights matrix is then given by

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{11} & \cdots & \mathbf{W}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{W}_{n1} & & \mathbf{W}_{nn} \end{pmatrix}, \tag{9}$$

which is an $N \times N$ block matrix, in which the blocks $\mathbf{W}_{ij}$ are $n \times n$ matrices, $N = n^2$.

We make the following assumption throughout:

Assumption 1 *The weights matrix* $\mathbf{W}$ *is an* $N \times N$ *(strictly) lower triangular matrix.*

The implication of Assumption 1 is that the process generating the data is assumed to be unilateral and of the type discussed in Section 2. The main reason for making this assumption is that the stationarity condition for spatial unilateral AR processes is well known and can easily be checked (see Yao and Brockwell 2006 for extension to causal SAR processes defined on the half plane). This is not the case for more general SAR processes. Extensions to nonunilateral and noncausal processes are briefly discussed in Section 5.

We now consider two examples in order to show how a SAR process can be represented by the SAR model (8) by a particular choice of the weights matrix $\mathbf{W}$.

Example 2 *Suppose that data are generated by the SAR process in (7). Then in (9) the blocks of matrices* $\mathbf{W}_{ii}$ *on the main diagonal are* $n \times n$

matrices with ones on the subdiagonal and all other elements are zeroes, and the blocks of matrices $\mathbf{W}_{ij}$ *on the subdiagonal are* $n \times n$ *identity matrices. A typical element of the vector* $\mathbf{y}$ *is*

$$y_{ij} = \rho(\frac{1}{2}y_{i-1,j} + \frac{1}{2}y_{i,j-1}),$$

for $i, j > 1$ *and after row standardisation of the weights matrix.*

Example 3 *Suppose that data are generated by the SAR process in (6). Then in (9) the blocks of matrices* $\mathbf{W}_{ii}$ *on the main diagonal are as in Example 2, and the blocks of matrices* $\mathbf{W}_{ij}$ *on the subdiagonal are* $n \times n$ *matrices with ones on both the main diagonal and the subdiagonal. A typical element of the vector* $\mathbf{y}$ *is*

$$y_{ij} = \rho(\frac{1}{3}y_{i-1,j} + \frac{1}{3}y_{i,j-1} + \frac{1}{3}y_{i-1,j-1})$$

for $i, j > 1$ *and after row standardisation of the weights matrix.*

The two examples illustrate how data generated by a unilateral SAR process can be represented by a SAR model.

The relation between the parameters in the unilateral SAR process and the SAR model is the following. The SAR model (8) is a reduced form model of the SAR process (3). The SAR model transforms the parameter space from $p = (p_1 + 1)(p_2 + 1) - 1$ dimensions to one dimension. In terms of the parameters of the SAR process, the parameter ρ is a linear combination with weights given by the non-zero elements in the weights matrix $\mathbf{W}$. The SAR model is a convenient statistical model in which the parameter ρ has the interpretation as a spatial autocorrelation coefficient.

It is worth noting that we cannot check the stationarity condition in the model (8), since we cannot from the parameter ρ recover the parameters of the SAR process. Anselin (1980) has shown that a necessary condition for stationarity in the SAR model is $1/\lambda_{\min} < \rho < 1/\lambda_{\max}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues of the weights matrix $\mathbf{W}$, respectively. This is a condition about the weights matrix, which is almost always satisfied. But the stationarity condition for the SAR process depends on the autoregressive parameters and cannot be checked based on the weights matrix. For any given weights matrix $\mathbf{W}$, there are points in the parameter space both in the stationary and non-stationary regions. But usually we do not care about the autoregressive parameters in the data generating process and as long as the process is stationary, we can estimate the SAR model (8) and give the parameter ρ its usual interpretation. Problems occur at the boundary of the parameter space.

We close this section by briefly remarking on issues of estimation and inference in the SAR model. The model can be solved in terms of $\boldsymbol{\varepsilon}$ as follows:

$$\mathbf{y} = (\mathbf{I}_N - \rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}, \tag{10}$$

provided that $(\mathbf{I}_N - \rho\mathbf{W})$ is non-singular. Hence, it is seen that the errors of the reduced form model are not IID. The model can be estimated by maximum likelihood by maximising the log likelihood function

$$\ln L = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\sigma^2 + \ln|\mathbf{I}_N - \rho\mathbf{W}| - \frac{1}{2}\mathbf{y}'(\mathbf{I}_N - \rho\mathbf{W})'(\mathbf{I}_N - \rho\mathbf{W})\mathbf{y}.$$

Under standard regularity conditions the ML estimator is consistent and asymptotically normal. For details, see Anselin (2003) and the references cited therein.

4 PROPERTIES OF THE ESTIMATOR

4.1 Large Sample Properties

Consider the SAR process (3). Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{(p_1+1)(p_2+1)-1})' = (\alpha_{10}, \ldots, \alpha_{p_1p_2})'$. We assume that $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta} \subset \mathbb{R}^{(p_1+1)(p_2+1)-1}$ is the parameter space. Following Yao and Brockwell (2006, p. 408), we assume the following condition holds:

(C1) The parameter space $\boldsymbol{\Theta}$ is a compact set containing the true value $\boldsymbol{\theta}_0$ as an interior point. Further, for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, condition (5) holds.

Let $\rho_0 = \mathbf{w}'\boldsymbol{\theta}_0$. The following proposition can be deduced from Theorem 1 in Yao and Brockwell (2006).

Proposition 4 *Let $\{\varepsilon_{ij}\} \sim IID(0, \sigma^2)$ and condition (C1) holds. Then as both N_1 and $N_2 \to \infty$, $\widehat{\rho} \xrightarrow{P} \rho_0$.*

Proof. The proof of Theorem 1 in Yao and Brockwell gives that $\widehat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}_0$ and by invariance of the maximum likelihood estimator $\widehat{\rho} = \mathbf{w}'\widehat{\boldsymbol{\theta}} \xrightarrow{P} \mathbf{w}'\boldsymbol{\theta}_0$. ■

The parameter ρ can be interpreted as a weighted average of the autoregressive parameters, with the weights given by the non-zero elements of the weights matrix. The result in Proposition 1 then says that the estimator $\widehat{\rho}$ consistently estimates a linear combination $\mathbf{w}'\boldsymbol{\theta}$ of the parameters $\boldsymbol{\theta}$, with the weights w_{ij} given by the non-zero elements of the spatial weights matrix $\mathbf{W}$. In the next section we investigate by simulation the bias of $\widehat{\rho}$ as an estimator of ρ.

Yao and Brockwell (2006) prove asymptotic normality of the estimator $\widehat{\boldsymbol{\theta}}$ under the additional condition (C2).

(C2) One of the following three conditions holds:

(i) $N_1 \to \infty$ and N_1/N_2 has a limit $d \in (0, \infty)$,

(ii) $N_2 \to \infty$ and $N_1/N_2 \to \infty$,

(iii) $N_1 \to \infty$ and $N_1/N_2 \to 0$.

The following proposition is a corollary to Theorem 2 in Yao and Brockwell (2006).

Proposition 5 *Let* $\{\varepsilon_{ij}\} \sim IID(0, \sigma^2)$ *and conditions (C1) and (C2) hold. Then* $N^{1/2}(\widehat{\rho} - \rho_0) \xrightarrow{D} N(0, \mathbf{w}'\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\mathbf{w})$.

Proof. The proof of Theorem 2 in Yao and Brockwell gives that $N^{1/2}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{D} \mathrm{N}(\boldsymbol{\theta}_0, \boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$. The final step in the proof follows from the Cramér–Wold device. ■

4.2 Finite Sample Properties

In this section we investigate by simulation the finite sample properties of the estimator $\widehat{\rho}$ in the SAR model. Data are generated from the SAR process (7) and the SAR model (8) is estimated with the weights matrix $\mathbf{W}_2$. The weights matrix $\mathbf{W}_2$ is given by (9), where the blocks of matrices $\mathbf{W}_{ii}$ on the main diagonal are $n \times n$ matrices with ones on the subdiagonal and all other elements are zeroes, and the blocks of matrices $\mathbf{W}_{ij}$ on the subdiagonal are $n \times n$ identity matrices. All other blocks are zero matrices. After row-standardisation of $\mathbf{W}_2$, the weights matrix has two non-zero elements in each row, except on the edges. The weights matrix $\mathbf{W}_2$ is correctly specified, since it captures the two-dimensional unilateral structure in the data generating process. Table 1 reports the mean of $\widehat{\rho}$ and Table 2 the mean of the bias of $\widehat{\rho}$ as an estimator of ρ. The number of observations are $N_1 = N_2 = 50$, so that $N = N_1 N_2 = 2500$. The number of replications is 1000. Note that the parameter space is symmetric in α_1 and α_2, so the results above and below the diagonal are from a single experiment. The bias is defined as $\text{bias}(\widehat{\rho}) = \widehat{\rho} - 0.5(\alpha_1 + \alpha_2)$. The bias of $\widehat{\rho}$ is small and the estimator is nearly unbiased, except at the boundary of the parameter space. For example, when $\alpha_1 = \alpha_2 = 0.2$ the bias is less than 0.000. For small values of $\alpha_1 + \alpha_2$ the bias is negative and for large values positive. For large values of $\alpha_1 + \alpha_2$ the bias becomes very large, as the parameters approach the boundary of the parameter space. At the boundary of the parameter space $\widehat{\rho}$ tends to 1. When $|\alpha_1| + |\alpha_2| > 1$, $\widehat{\rho}$ diverges.

We now consider the cases when the weights matrix is under- or over-specified. In the case of an underspecified weights matrix, the SAR model is estimated with the weights matrix $\mathbf{W}_1$. The weights matrix $\mathbf{W}_1$ is again given by (9), where the blocks of matrices $\mathbf{W}_{ii}$ on the main diagonal are $n \times n$ matrices with ones on the subdiagonal and all other elements are zeroes. All other blocks are zero matrices. The weights matrix $\mathbf{W}_1$ has only one non-zero element in each row, except on the edges. The weights matrix $\mathbf{W}_1$ is underspecified, since it represents only one-dimension of the two-dimensional unilateral structure in the data generating process. The bias is defined as $\text{bias}(\widehat{\rho}) = \widehat{\rho} - 0.5(\alpha_1 + \alpha_2)$. The bias is negative for all values of α_1 and α_2, except at the boundary of the parameter space. Underspecifying the weights matrix results in an underestimated ρ. Finally, we consider the case of an overspecified weights matrix. The SAR model is estimated with the weights matrix $\mathbf{W}_4$. The weights matrix $\mathbf{W}_4$ is chosen to be no longer lower triangular. The weights matrix $\mathbf{W}_4$ is again given by (9), where the blocks of matrices $\mathbf{W}_{ii}$ on the main diagonal are $n \times n$ matrices with ones both on the subdiagonal and superdiagonal, and all other elements are zeroes. The blocks of matrices $\mathbf{W}_{ij}$ on the subdiagonal and superdiagonal are $n \times n$ identity matrices. All other blocks are zero matrices. After row-standardisation of $\mathbf{W}_4$, the weights matrix has four non-zero elements in each row, except on the edges. The weights matrix $\mathbf{W}_4$ is overspecified. It represents a two-dimensional bilateral structure, but the data generating process is unilateral. The estimated SAR model is therefore noncausal. The bias is defined as $\text{bias}(\widehat{\rho}) = \widehat{\rho} - 0.5(\alpha_1 + \alpha_2)$. The bias is positive for all values of α_1 and α_2. Overspecifying the weights matrix results in an overestimated ρ.

Returning to the problem of checking the stationarity condition in the SAR model, we see that an estimated ρ close to 1 is a clear indication that the estimated model is nonstationary. It is obvious that inference in the SAR model is linked with the problem of determining the weights matrix. In

Table 1: The mean of $\hat{\rho}$ in the SAR model (8). The number of observations is $N = 2500$ and the number of replications is 1000.

α_1/α_2	0	0.05	0.10	0.20	0.30	0.40	0.45	0.50
				Weights matrix $\mathbf{W}_1$				
0	0.000	0.026	0.050	0.100	0.154	0.213	0.241	0.277
0.05	−0.000	0.026	0.050	0.100	0.155	0.213	0.243	0.278
0.10	−0.000	0.025	0.051	0.101	0.156	0.215	0.245	0.281
0.20	−0.000	0.026	0.052	0.104	0.162	0.224	0.257	0.295
0.30	−0.000	0.027	0.054	0.109	0.173	0.240	0.281	0.324
0.40	−0.001	0.029	0.058	0.119	0.190	0.272	0.323	0.386
0.45	−0.001	0.031	0.061	0.126	0.203	0.297	0.359	0.451
0.50	−0.001	0.033	0.065	0.135	0.219	0.333	0.423	0.597
				Weights matrix $\mathbf{W}_2$				
0	0.000	0.024	0.047	0.096	0.145	0.197	0.222	0.251
0.05	0.024	0.047	0.072	0.121	0.172	0.225	0.255	0.285
0.10	0.047	0.072	0.096	0.147	0.199	0.257	0.288	0.319
0.20	0.096	0.121	0.147	0.200	0.260	0.324	0.360	0.401
0.30	0.145	0.172	0.199	0.260	0.326	0.405	0.450	0.502
0.40	0.197	0.225	0.257	0.324	0.405	0.503	0.559	0.634
0.45	0.222	0.255	0.288	0.360	0.450	0.559	0.634	0.731
0.50	0.251	0.285	0.319	0.401	0.502	0.634	0.731	0.872
				Weights matrix $\mathbf{W}_4$				
0	−0.000	0.048	0.096	0.192	0.286	0.377	0.424	0.469
0.05	0.048	0.095	0.144	0.239	0.331	0.425	0.470	0.514
0.10	0.096	0.144	0.192	0.286	0.378	0.470	0.514	0.553
0.20	0.192	0.239	0.286	0.378	0.470	0.554	0.596	0.640
0.30	0.286	0.331	0.378	0.470	0.553	0.640	0.684	0.725
0.40	0.377	0.425	0.470	0.554	0.640	0.724	0.766	0.813
0.45	0.424	0.470	0.514	0.596	0.684	0.766	0.812	0.860
0.50	0.469	0.514	0.553	0.640	0.725	0.813	0.860	0.923

Table 2: The mean of the bias of $\hat{\rho}$ in the SAR model (8). The number of observations is $N = 2500$ and the number of replications is 1000.

α_1/α_2	0	0.05	0.10	0.20	0.30	0.40	0.45	0.50
				Weights matrix $\mathbf{W}_1$				
0	0.000	0.000	0.000	−0.000	0.004	0.013	0.016	0.027
0.05	−0.025	−0.025	−0.025	−0.025	−0.020	−0.012	−0.008	0.003
0.10	−0.050	−0.050	−0.049	−0.049	−0.044	−0.035	−0.030	−0.019
0.20	−0.100	−0.099	−0.099	−0.096	−0.088	−0.076	−0.068	−0.055
0.30	−0.150	−0.148	−0.146	−0.141	−0.128	−0.110	−0.095	−0.076
0.40	−0.201	−0.196	−0.192	−0.181	−0.160	−0.128	−0.103	−0.064
0.45	−0.226	−0.219	−0.214	−0.199	−0.172	−0.128	−0.091	−0.024
0.50	−0.251	−0.242	−0.235	−0.215	−0.181	−0.117	−0.052	0.097
				Weights matrix $\mathbf{W}_2$				
0	0.000	−0.001	−0.003	−0.004	−0.005	−0.003	−0.003	0.000
0.05	−0.001	−0.003	−0.003	−0.005	−0.030	−0.000	0.005	0.010
0.10	−0.003	−0.003	−0.004	−0.004	−0.001	0.007	0.013	0.019
0.20	−0.004	−0.005	−0.004	0.000	0.010	0.024	0.035	0.051
0.30	−0.005	−0.003	−0.001	0.010	0.026	0.055	0.075	0.102
0.40	−0.003	−0.000	0.007	0.024	0.055	0.103	0.134	0.184
0.45	−0.003	0.005	0.013	0.035	0.075	0.134	0.184	0.256
0.50	0.000	0.010	0.019	0.051	0.102	0.184	0.256	0.372
				Weights matrix $\mathbf{W}_4$				
0	−0.000	0.023	0.046	0.092	0.136	0.177	0.199	0.219
0.05	0.023	0.045	0.069	0.114	0.156	0.200	0.220	0.239
0.10	0.046	0.069	0.092	0.136	0.178	0.220	0.239	0.253
0.20	0.092	0.114	0.136	0.178	0.220	0.254	0.271	0.290
0.30	0.136	0.156	0.178	0.220	0.253	0.290	0.309	0.325
0.40	0.177	0.200	0.220	0.254	0.290	0.324	0.341	0.363
0.45	0.199	0.220	0.239	0.271	0.309	0.341	0.362	0.385
0.50	0.219	0.239	0.253	0.290	0.325	0.363	0.385	0.423

another paper (Gerkman and Ahlgren 2007) we look at selection procedures for the weights matrix.

4.3 Spatial Error Models

Spatial dependence can be incorporated in linear regression models through a spatially lagged dependent variable or spatial dependence in the errors. The SAR model with spatially lagged explanatory variables is given by

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{X} = [x_{ij1}, \ldots, x_{ijk}]$ is an $N \times k$ matrix of explanatory variables, $\boldsymbol{\beta}$ is a $k \times 1$ parameter vector, and $\mathbf{y}$, $\mathbf{W}$ and $\boldsymbol{\varepsilon}$ are as before. The spatial error model (SEM) can be written as

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \\ \mathbf{u} &= \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \end{aligned} \tag{11}$$

where $\mathbf{X} = [x_{ij1}, \ldots, x_{ijk}]$ is an $N \times k$ matrix of explanatory variables, $\boldsymbol{\beta}$ is a $k \times 1$ parameter vector, $\mathbf{u} = [u_{ij}]$ is an $N \times 1$ vector of errors, λ is a scalar parameter, and $\mathbf{y}$, $\mathbf{W}$ and $\boldsymbol{\varepsilon}$ are as before. The spatial error model is commonly used to correct for the potentially biasing influence of spatial autocorrelation present in spatial data. The LS estimator of $\boldsymbol{\beta}$ is unbiased and consistent in the SEM, but is not efficient and the standard errors will be biased.

We study by simulation the finite sample properties of the estimators of the spatial autoregressive parameter λ and the parameter vector $\boldsymbol{\beta}$ in the SEM. Data are generated from the following SEM DGP:

$$\begin{aligned} y_{ij} &= \beta x_{ij} + u_{ij}, \\ u_{ij} &= \alpha_1 u_{i-1,j} + \alpha_2 u_{i,j-1} + \varepsilon_{ij}. \end{aligned} \tag{12}$$

The exogenous explanatory variable x_{ij} is generated as $x_{ij} \sim \text{NID}(0,1)$ with $\beta = 1$. In (12) the errors u_{ij} are generated by a unilateral spatial AR process as in (7) and $\varepsilon_{ij} \sim \text{NID}(0, 12.5^2)$. The number of observations are $N_1 = N_2 = 50$, so that $N = N_1N_2 = 2500$. With these parameter values the t-statistic of $\beta = 0$ when $\alpha_1 = \alpha_2 = 0$ is 4 in the population. In the simulations we set $\alpha_1 = \alpha_2 = \alpha$, so there is a single spatial parameter α in the simulation experiments. The number of replications is 1000. The SEM

(11) is first estimated with the weights matrix $\mathbf{W}_2$, which is assumed to capture the spatial structure in the data correctly, since it has two non-zero elements in each row. The SEM is also estimated using the underspecified weights matrix $\mathbf{W}_1$ and the overspecified weights matrix $\mathbf{W}_4$. The weights matrix $\mathbf{W}_0$ refers to the model without spatial errors and then the maximum likelihood estimator of β is equivalent to the LS estimator. Table 3 reports the mean of $\hat{\lambda}$, the rejection probability of the null hypothesis $\lambda = 0$ at the 5% significance level, the mean of $\hat{\beta}$, the mean of the bias of $\hat{\beta}$ and the rejection probability of the null hypothesis $\beta = 0$ at the 5% significance level. The results show that as before, the estimate of λ depends on the weights matrix. The test of $\lambda = 0$ has the correct size about 5% when $\alpha = 0$ and high power, which tends to one as $\alpha \geq 0.10$. For estimating $\boldsymbol{\beta}$ in the SEM, the choice of weights matrix does not seem to be all that important, since $\widehat{\beta}$ is nearly unbiased, independent of the choice of weights matrix. For large values of α, the t-value of $\beta = 0$ is underestimated and the rejection probability of $\beta = 0$ decreases in the model without spatial errors.

5 EXTENSIONS OF THE MODEL

Here we briefly discuss some extensions of the unilateral SAR process and the SAR model. There are several extensions of the unilateral SAR model that can be pursued. First, the unilateral SAR process on the quarter plane can be extended to the half plane (Yao and Brockwell 2006). Second, the model can be extended to nonunilateral processes (see Tjøstheim 1981 for a discussion). For example, in spatial econometrics two commonly used specifications of the weights matrix are 'rook' and 'queen' contiguities, which are based on weights matrices that are not lower triangular (see Anselin 1988, p. 22). The stationarity conditions for the SAR model are then not known and the model is noncausal. Since all SAR models are approximations, the question is under what conditions can spatial data be approximated by unilateral SAR models (see Tjøstheim 1981 for more on this). In the application to house prices data in the next section we will entertain both unilateral and nonunilateral specifications of the weights matrix. Finally, to model data over both space and time, extensions to spatio-temporal modelling can be considered along the lines in Yao and Brockwell (2006).

Table 3: The mean of $\hat{\lambda}$, the rejection probability of the null hypothesis $\lambda = 0$ at the 5% significance level, the mean of $\hat{\beta}$, the mean of the bias of $\hat{\beta}$ and the rejection probability of the null hypothesis $\beta = 0$ at the 5% level in the SEM (11) when data are generated by the SEM DGP (12). The number of observations is $N = 2500$ and the number of replications is 1000.

$\alpha_1 = \alpha_2$	$\widehat{\lambda}$	$\lambda = 0$	$\widehat{\beta}$	bias($\widehat{\beta}$)	$\beta = 0$
		Weights matrix $\mathbf{W}_0$			
0			1.006	0.006	0.981
0.05			1.006	0.006	0.983
0.10			1.006	0.006	0.984
0.20			1.006	0.006	0.974
0.30			1.007	0.007	0.954
0.40			1.006	0.006	0.891
0.45			1.005	0.005	0.754
0.50			0.993	−0.007	0.382
		Weights matrix $\mathbf{W}_1$			
0	−0.000	0.051	1.007	0.007	0.981
0.05	0.050	0.704	1.006	0.006	0.982
0.10	0.101	1.000	1.006	0.006	0.983
0.20	0.209	1.000	1.005	0.005	0.983
0.30	0.333	1.000	1.004	0.004	0.983
0.40	0.497	1.000	1.002	0.002	0.986
0.45	0.619	1.000	1.000	−0.000	0.984
0.50	0.844	1.000	0.997	−0.003	0.982
		Weights matrix $\mathbf{W}_2$			
0	0.001	0.055	1.007	0.007	0.982
0.05	0.097	0.929	1.006	0.006	0.979
0.10	0.193	1.000	1.006	0.006	0.979
0.20	0.387	1.000	1.005	0.005	0.985
0.30	0.582	1.000	1.005	0.005	0.989
0.40	0.782	1.000	1.004	0.004	0.995
0.45	0.885	1.000	1.003	0.003	0.998
0.50	0.987	1.000	1.003	0.003	0.999
		Weights matrix $\mathbf{W}_4$			
0	−0.000	0.412	1.007	0.007	0.982
0.05	0.097	0.963	1.006	0.006	0.980
0.10	0.195	1.000	1.006	0.006	0.981
0.20	0.380	1.000	1.005	0.005	0.982
0.30	0.556	1.000	1.004	0.004	0.992
0.40	0.727	1.000	1.003	0.003	0.996
0.45	0.816	1.000	1.003	0.003	0.997
0.50	0.927	1.000	1.003	0.003	0.998

6 APPLICATION TO HOUSE PRICES

In this section we provide an application to house prices. The data consist of 1377 transactions of single-family houses between January 2000 and May 2001 in the county of Stockholm, Sweden. The data were analysed by Wilhelmsson (2002) and include the selling price, spatial coordinates and in addition information about the size of the house in square metres as well as other characteristics. See Wilhelmsson (2002) for a detailed description of the data. The variables are defined as follows: P is price in SEK, LA is living area in square metres, OA is other area in square metres, A is age in years, $A60$ is a dummy variable taking the value 1 if $A \geq 60$ and 0 otherwise, SV is a dummy variable for sea view, LS is lot size in square metres, $Q1$ is a quality index, D is distance to the city centre i metres, $Q22000, \ldots, Q22001$ are dummy variables for quarter. The estimated models also include 12 dummies for parish (not reported). All variables are in logs.

We estimate SEM models (11) with the weights matrices given by $\mathbf{W}_0$, $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{W}_4$. The weights matrices are defined as in Section 4.3 and are based on the closest neighbours. We find that the magnitude of the estimated spatial autocorrelation coefficient λ in the SEM model crucially depends on the choice of weights matrix. This result is not a surprise, given the discussion of the interpretation of the spatial autocorrelation coefficient in Section 3 and the simulation results in Section 4. If we compare the models with one and four neighbours, we find that in the model with one neighbour $\widehat{\lambda} = 0.125$ and in the model with four neighbours $\widehat{\lambda} = 0.278$, i.e., more than twice as large. It would of course be useful to know which one of the weights matrices $\mathbf{W}_0$, $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{W}_4$ should be preferred and we will return to this problem in another paper (Gerkman and Ahlgren 2007). The estimated coefficients on the explanatory variables do not change much with different weights matrices, which is in line with the findings in Section 4.3. Note that the t-values are large, so they are of no help in selecting explanatory variables in the SEM models.

A comment on Wilhelmsson (2002) is in order. As an alternative specification of the weights matrix, Wilhelmsson uses the inverse of the distance between the observations $(1/d_{ij})$ as elements in the weights matrix. In this case an estimate of λ close to 1 is obtained. The reported (restricted?) estimate is $\widehat{\lambda} = 0.998$. The estimated model is most likely non-stationary, which can be explained by the well known fact that the harmonic series $\sum_{d=1}^{n} 1/d$ diverges, as $n \to \infty$. This seems to be a rare example of a SEM model with

Table 4: Comparison of SEM models for house prices in the county of Stockholm, Sweden with different weights matrices. The first column for each weights matrix reports the parameter estimates and the second column the asymptotic t-ratios.

	$\mathbf{W}_0$		$\mathbf{W}_1$		$\mathbf{W}_2$		$\mathbf{W}_4$	
LA	0.508	21.701	0.485	20.863	0.478	20.728	0.474	20.055
OA	0.020	3.249	0.018	3.083	0.017	2.882	0.015	2.599
A	−0.047	−2.400	−0.074	−4.252	−0.075	−4.330	−0.072	−4.149
$A60$	0.118	6.384	0.106	5.785	0.102	5.657	0.102	5.524
SV	0.335	5.300	0.267	4.251	0.253	4.190	0.245	3.956
LS	0.156	6.255	0.142	7.581	0.146	8.344	0.145	7.628
QI	0.236	6.267	0.207	7.020	0.204	7.178	0.207	7.026
D	−0.304	−6.839	−0.447	−105.055	−0.444	−100.808	−0.444	−99.369
$Q22000$	0.112	5.130	0.106	5.065	0.101	4.862	0.100	4.835
$Q32000$	0.151	6.789	0.148	6.907	0.145	6.824	0.147	6.955
$Q42000$	0.229	10.505	0.224	10.639	0.224	10.753	0.226	10.963
$Q12001$	0.221	9.724	0.214	9.733	0.216	9.896	0.222	10.273
$Q22001$	0.290	9.694	0.291	10.093	0.292	10.182	0.292	10.206
λ			0.125	9.536	0.193	10.785	0.278	21.180
$\log L$			502.086		506.190		513.094	
R^2			0.683		0.686		0.689	
Moran I			0.225		0.173		0.145	
I stat			6.925		7.440		8.829	

a weights matrix that is nonstationary. On the other hand, if the measure of distance is the squared inverse of the distance $(1/d_{ij}^2)$, the estimate is reduced to $\widehat{\lambda} = 0.360$. The estimated model appears to be stationary, which again follows from the simple fact that the series $\sum_{d=1}^{n} 1/d^2$ converges, as $n \to \infty$.

7 CONCLUSIONS

In this paper we have studied inference in spatial econometric models. First, we give conditions for causality and stationarity in spatial autoregressive models, based on the theory of unilateral spatial autoregressive processes. Second, we investigate the properties of the maximum likelihood estimator of the spatial autocorrelation coefficient. In the stationary case the estimator is shown to be consistent and asymptotically normal. In the nonstationary case the estimator is inconsistent. Third, we investigate the impact of the weights matrix on the estimated spatial autocorrelation coefficient. We find that the estimated spatial autocorrelation coefficient crucially depends on the weights matrix. Thus, depending on the choice of weights matrix a very different picture of the magnitude of spatial autocorrelation is obtained. As an illustration, an application to house prices is provided.

References

[1] Anselin, L. (1980), *Estimation Methods for Spatial Autoregressive Structures*, Regional Science Dissertation and Monograph Series 8, Cornell University, Ithaca, NY.

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

[3] Anselin, L. (2003), 'Spatial Econometrics', in Baltagi, B. H. (ed.), *A Companion to Theoretical Econometrics*, Blackwell Publishing, Oxford, 310–330.

[4] Baran, S., Pap, G. and Zuijlen, M. C. A. van (2005), 'Asymptotic Inference for Unit Roots in Spatial Autoregression', Proceedings of the 25th European Meeting of Statisticians, Oslo, 254.

[5] Basu, S. and Reinsel, G. C. (1993), 'Properties of the Spatial Unilateral First-Order ARMA Model', *Advances in Applied Probability*, 25, 631–648.

[6] Besag, J. E. (1972), 'On the Correlation Structure of some Two-Dimensional Stationary Processes', *Biometrika*, 59, 43-48.

[7] Gerkman, L. and Ahlgren, N. (2006), Selection Procedures for the Weights Matrix in Spatial Econometric Models, Manuscript, Swedish School of Economics.

[8] Tjøstheim, D. (1978), 'Statistical Spatial Series Modelling', *Advances in Applied Probability*, 10, 130–154.

[9] Tjøstheim, D. (1981), 'Autoregressive Modeling and Spectral Analysis of Array Data in the Plane', *IEEE Transactions on Geoscience and Remote Sensing*, 19, 15-24.

[10] Tjøstheim, D. (1983), 'Statistical Spatial Series Modelling II: Some further Results in Unilateral Processes', *Advances in Applied Probability*, 15, 562–684.

[11] Whittle, P. (1954), 'On Stationary Processes in the Plane', *Biometrika*, 41, 434-449.

[12] Wilhelmsson, M. (2002), 'Spatial Models in Real Estate Economics, *Housing, Theory and Society*, 19, 92-101.

[13] Yao, Q. and Brockwell, P. J. (2006), 'Gaussian Maximum Likelihood Estimation for ARMA Models II: Spatial Processes', *Bernoulli*, 13, 403–429.