

Folmer, Henk; Oud, Johan

Conference Paper

A Structural Equation Approach to Spatial Dependence Models

46th Congress of the European Regional Science Association: "Enlargement, Southern Europe and the Mediterranean", August 30th - September 3rd, 2006, Volos, Greece

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Folmer, Henk; Oud, Johan (2006) : A Structural Equation Approach to Spatial Dependence Models, 46th Congress of the European Regional Science Association: "Enlargement, Southern Europe and the Mediterranean", August 30th - September 3rd, 2006, Volos, Greece, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/118164>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**A STRUCTURAL EQUATION APPROACH TO SPATIAL
DEPENDENCE MODELS**

Johan Oud ¹

Henk Folmer ²

¹Behavioural Science Institute, Radboud University Nijmegen,
P.O. Box 9104, NL-6500 HE Nijmegen, The Netherlands, E-mail: j.oud@pwo.ru.nl

²Department of Social Sciences, Wageningen University, P.O. Box 8130,
NL-6700 EW Wageningen, The Netherlands, E-mail: Henk.Folmer@WUR.NL

Abstract

In this paper we propose a Structural Equations Model (SEM) approach to spatial dependence models. Latent variables are used to represent spatial spill-over effects in the structural model of which the observed spatially lagged variables are indicators. This approach allows for more information and modeling flexibility than the representation of spatial spill-over effects in terms of Wy or Wx . Furthermore, we propose a Full Information Maximum Likelihood (FIML) estimator as an alternative to the estimators commonly used, notably the iterative and two-stage estimators for the error and lag model, respectively. We also show that the estimation procedures included in the software packages Mx and LISREL 8 to estimate SEMs can be applied in a straightforward way to estimate spatial dependence models in a standard fashion.

1. Introduction

During the past decades several estimation procedures of linear models with spatial dependence have been developed. Well-known and frequently applied are the iterative and two-stage estimators for the error and the lag model. Particularly, Anselin (1988) proposes an iterative two-stage procedure to maximize the log-likelihood of the spatial error model. From the initial OLS estimator of the regression coefficients the residuals are calculated. Given the residuals, the spatial correlation parameter is estimated from maximization of the concentrated log-likelihood. Given the spatial parameter estimated, feasible GLS is applied to obtain new estimates of the regression coefficients and to compute a new set of residuals. When a convergence criterion is met, the lastly obtained estimates are taken as the final ones and the variance of the disturbance is estimated. Otherwise, the above procedure is repeated until the convergence criterion is met (see Anselin and Hudak (1992) for details).

For the spatial lag model Anselin (1988) proposes a non-iterative two-stage procedure. The vectors of regression coefficients obtained from OLS of the dependent variable and of the spatially lagged dependent variable on the exogenous variables, respectively, are used to obtain the respective sets of residuals. Given the sets of residuals, the spatial autocorrelation parameter is obtained from maximization of the concentrated log-likelihood.

Given the spatial autocorrelation parameter, the vector of regression coefficients of the spatial lag model and the variance of the disturbance term are obtained (see Anselin and Hudak (1992) for details).

Typical for the class of estimation procedures of spatial dependence models presently in use is that they are restricted to models that contain directly observable variables only. The purpose of this paper is to introduce the class of Structural Equations Models (SEM) and corresponding estimation procedures that allow within one and the same model framework the presence of both latent and observable variables. Latent variables (also denoted theoretical constructs) refer to those phenomena that are supposed to exist but cannot be observed directly. Well-known examples of latent variables are utility, socio-economic status, regional welfare. Directly observable variables on the other hand possess direct empirical meanings derived from experience. Latent variables can only be observed and measured by means of observable variables. For instance, the latent variable socio-economic status is observed and measured by such observable variables as income, education, profession, position in a social network, etc. In a similar vein, regional welfare is measured by regional observable variables such as per capita GDP, income distribution, employment opportunities, features of the housing market, health indicators, indicators of environmental quality, etc. The simultaneous use of both latent and observable

variables in an empirical analysis has amongst others the advantages that latent variables are give empirical meanings by means of operational definitions; that a closer correspondence between theory and empirics can be obtained; that measurement errors can be accounted for, and that the impacts of multicollinearity can be mitigated. (See amongst others Blalock and Blalock (1971) and Folmer (1986) and the references therein for further details.)

The class of Structural Equations Models (SEM) that we are considering in this paper makes it possible to simultaneously estimate theoretical statements (which contain latent variables only) and correspondence statements (which contain both latent and observable variables). Particularly, a SEM is made up of two related sub-models:

- A latent variables measurement model which represents the relationships between the latent variables and their observable indicators.

- A structural model which represents the relationships between the latent variables.

An immediate consequence of the presence of both latent and observable variables within one model framework is that it allows for an alternative representation of spatial dependence in the sense that the spatially lagged variables Wy (spatial lag model) or Wx (spatial cross model) can

be represented by means of latent variables. We will first generalize the spatial lag to models containing a latent dependent variable. Next, the conventional way of representing the spatial model by means of a weight matrix W describing the spatial arrangement of the units of observation is replaced by latent variables representing spatial dependence of which the observed values of the neighbouring spatial units are indicators. As will be shown, the latent variable approach to modelling spatial dependence allows a more informative and more flexible presentation of spatial dependence than the conventional approach by means of a weight matrix that is given a priori.

In this paper we furthermore present a Full Information Maximum Likelihood (FIML) alternative estimator to the iterative and two-stage procedure. That is, the coefficients of interest are estimated in a simultaneous equations model framework such that the nature of spatial dependence, i.e. the (two-way) interaction between the dependent variable and its spatial lags³, is adequately taken into account.

This paper is organized as follows. In section 2 we briefly introduce the class of SEMs. In section 3 we specify the standard lag and error model as SEMs and estimate them for Anselin's (1988) Columbus, Ohio, crime data set, applying the standard software packages Mx and LISREL 8 in a bid to show that standard spatial dependence models can be routinely

estimated by SEM software. In section 4 we present the lag model for a latent dependent variable. In section 5 we present a SEM representation of the spatial lag model such that in the structural model Wy is replaced by a latent variable (spatial spill-over) and the measurement model represents the relationships between the latent variable and the lagged observed dependent variables in neighbouring units of observation. The models in sections 4 and 5 are applied to the Columbus, Ohio, crime data set again. Section 5 concludes the paper.

2. The Structural Equation Model (SEM)

A SEM in general form reads:

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta} \quad \text{with} \quad \text{cov}(\boldsymbol{\zeta}) = \boldsymbol{\Psi} , \quad (1)$$

$$\mathbf{y} = \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad \text{with} \quad \text{cov}(\boldsymbol{\varepsilon}) = \boldsymbol{\Theta} . \quad (2)$$

The model consists of two equations. In the latent structural equation (1) vector $\boldsymbol{\eta}$ contains the k latent variables, \mathbf{B} specifies the structural relationships among the latent variables, and $\boldsymbol{\Psi}$ is the covariance matrix of the vector of errors $\boldsymbol{\zeta}$. Moreover, $\boldsymbol{\Psi}$ encompasses the covariance matrix of the exogenous variables, contained in $\boldsymbol{\eta}$. In the measurement equation (2) the vector \mathbf{y} contains the p observed variables, the $p \times k$ -matrix $\mathbf{\Lambda}$ specifies the loadings or regression coefficients of the observed variables on the latent variables, and $\boldsymbol{\Theta}$ is the measurement error covariance matrix.

The measurement errors in $\boldsymbol{\varepsilon}$ are assumed to be uncorrelated with the latent variables in $\boldsymbol{\eta}$ as well as with the structural errors in $\boldsymbol{\zeta}$.

Although SEM originated in the field of confirmatory factor analysis, it encompasses a wide variety of classes of models, such as first and second order factor analysis models, structural equation models for directly observable variables and various types of regression models. For instance, if in (2) $\boldsymbol{\Lambda} = \mathbf{I}$ and $\boldsymbol{\varepsilon} = \mathbf{0}$ the structural equation model for directly observable variables results.

Instead of model (1)-(2) one often encounters a more elaborate version of the SEM model in which endogenous and exogenous variables are put in different latent vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ and observed vectors \mathbf{y} and \mathbf{x} , respectively:

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \quad \text{with} \quad \text{cov}(\boldsymbol{\xi}) = \boldsymbol{\Phi}, \text{cov}(\boldsymbol{\zeta}) = \boldsymbol{\Psi} , \quad (3)$$

$$\mathbf{y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad \text{with} \quad \text{cov}(\boldsymbol{\varepsilon}) = \boldsymbol{\Theta}_\varepsilon , \quad (4)$$

$$\mathbf{x} = \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta} \quad \text{with} \quad \text{cov}(\boldsymbol{\delta}) = \boldsymbol{\Theta}_\delta . \quad (5)$$

However, the formulation in model (1)-(2) with all latent variables in one single vector $\boldsymbol{\eta}$ and all observed variables in another single vector \mathbf{y} , is in fact more flexible, making it possible, for instance, to directly specify correlations among endogenous measurement errors $\boldsymbol{\varepsilon}$ and exogenous measurement errors $\boldsymbol{\delta}$.

Several parameter estimation methods for SEMs have been developed including instrumental variables (IV), two-stage least squares (TSLS),

unweighted least squares (ULS), generalized least squares (GLS), fully weighted (WLS) and diagonally weighted least squares (DWLS), and maximum likelihood (ML). Many easily accessible software packages are available to estimate SEMs including Mx (Neale, Boker, Xie, & Maes, 2003) and LISREL 8 (Jöreskog, K.G., & Sörbom, D., 1996). These packages include procedures to check model identification, to evaluate the estimation results and to calculate indirect and total effects. Here we restrict ourselves to the ML method, which maximizes the loglikelihood function of the free elements in the parameter matrices \mathbf{B} , $\mathbf{\Psi}$, $\mathbf{\Lambda}$, and $\mathbf{\Theta}$ for given data \mathbf{Y} :

$$\ell(\boldsymbol{\theta}|\mathbf{Y}) = -\frac{N}{2} \ln |\boldsymbol{\Sigma}| - \frac{N}{2} \text{tr}(\mathbf{S}\boldsymbol{\Sigma}^{-1}) - \frac{pN}{2} \ln 2\pi . \quad (6)$$

$\boldsymbol{\theta}$ in (6) contains the free parameters in the matrices \mathbf{B} , $\mathbf{\Psi}$, $\mathbf{\Lambda}$, and $\mathbf{\Theta}$, $\mathbf{Y}_{N \times p}$ is the data matrix (N rows of independent replications of the p -variate vector \mathbf{y} , typically originating from a sample of randomly drawn subjects), $\boldsymbol{\Sigma}_{p \times p}$ is the model-implied covariance or moment matrix:

$$\boldsymbol{\Sigma} = \mathbf{\Lambda}(\mathbf{I} - \mathbf{B})^{-1} \mathbf{\Psi} (\mathbf{I} - \mathbf{B})^{-1} \mathbf{\Lambda}' + \mathbf{\Theta} , \quad (7)$$

which is a function $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ of $\boldsymbol{\theta}$. Finally, $\mathbf{S}_{p \times p} = \frac{1}{N} \mathbf{Y}'\mathbf{Y}$ is the sample covariance or moment matrix.

The ML-estimator $\hat{\boldsymbol{\theta}} = \text{argmax } \ell(\boldsymbol{\theta}|\mathbf{Y})$ chooses that value of $\boldsymbol{\theta}$ which maximizes $\ell(\boldsymbol{\theta}|\mathbf{Y})$. If the observed variables follow a multivariate normal distribution, maximization of $\ell(\boldsymbol{\theta}|\mathbf{Y})$ gives genuine maximum likelihood es-

timates. However, when the range of the variables is in principle $(-\infty, \infty)$ and second-order moments exist, the assumption of multivariate normality can be justified as a first working hypothesis on the basis of a central limit theorem or maximum entropy. The latter means that the normal distribution reflects the lack of knowledge about the distribution more completely than other distributions (Rao, 1965).

Application of ML under the assumption of normality whereas the distribution actually deviates from normality may also be defended on the basis of the fact that it usually leads to a reasonable fitting function and to estimators with acceptable properties for a wide class of distributions. However, in the case of deviation from normality the standard errors produced by LISREL 8 and most other SEM programs should be interpreted with caution. The same applies to various statistics for model fit judgement, especially χ^2 .

The ML fitting function can also be used without the assumption of normality. Under these circumstances the estimator is still consistent. However, the model fit judgement statistics are no longer valid. Similar observations apply to the other estimators in the LISREL 8 and other SEM programs.

Instead of maximizing the loglikelihood function in (6) standard soft-

ware for SEM analysis usually minimizes the fit function

$$F_{ML} = \ln |\boldsymbol{\Sigma}| + \text{tr}(\mathbf{S}\boldsymbol{\Sigma}^{-1}) - \ln |\mathbf{S}| - p \quad (8)$$

or $\chi^2 = (N - 1)F_{ML}$ with the same result. Because the data based \mathbf{S} is a constant, (6) and (8) relate linearly.

As an introduction to the next section, we consider the vector of exogenous variables $\boldsymbol{\xi}$ in (3) as fixed and observed, that is $\boldsymbol{\xi} = \mathbf{x}$ (no equation (5)). In that case the loglikelihood function (6) reduces to (Oud, 2004)

$$\begin{aligned} \ell(\boldsymbol{\theta}|\mathbf{Y}_0) = & \\ & -\frac{N}{2} \ln |\boldsymbol{\Sigma}_0| - \frac{1}{2} \sum_{i=1}^N (\mathbf{y}_{0i} - \boldsymbol{\mu}_{0i})' \boldsymbol{\Sigma}_0^{-1} (\mathbf{y}_{0i} - \boldsymbol{\mu}_{0i}) - \frac{p_0 N}{2} \ln 2\pi, \end{aligned} \quad (9)$$

where the subscript of \mathbf{Y}_0 and \mathbf{y}_0 indicates that no exogenous variables are included, p_0 is the number of variables in \mathbf{Y}_0 and \mathbf{y}_0 ,

$$\boldsymbol{\mu}_0 = E(\mathbf{y}_0) = \boldsymbol{\Lambda}_y (\mathbf{I} - \mathbf{B})^{-1} \boldsymbol{\Gamma} \mathbf{x}, \quad (10)$$

$$\boldsymbol{\Sigma}_0 = E[(\mathbf{y}_0 - \boldsymbol{\mu}_0)(\mathbf{y}_0 - \boldsymbol{\mu}_0)'] = \boldsymbol{\Lambda}_y \boldsymbol{\Psi} \boldsymbol{\Lambda}_y' + \boldsymbol{\Theta}_\varepsilon. \quad (11)$$

For a latent regression model (MIMIC or multiple-indicators-multiple-causes model, see Jöreskog, K.G., & Sörbom, D., 1996, p. 185-187), a further simplification applies: $\mathbf{B} = \mathbf{0}$, $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ in (10) and (11) become

$$\boldsymbol{\mu}_0 = \boldsymbol{\Lambda}_y \boldsymbol{\Gamma} \mathbf{x}, \quad (12)$$

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda}_y \boldsymbol{\Psi} \boldsymbol{\Lambda}_y' + \boldsymbol{\Theta}_\varepsilon. \quad (13)$$

3. SEM representation of spatial dependence

Spatial lag model in SEM

We consider the standard spatial lag and spatial error models. First the spatial lag model:

$$\tilde{\mathbf{y}} = \rho \mathbf{W} \tilde{\mathbf{y}} + \mathbf{X} \boldsymbol{\gamma} + \tilde{\boldsymbol{\epsilon}}, \quad (14)$$

where

$\tilde{\mathbf{y}}$ is the $N \times 1$ vector with observations on the dependent variable y ;

\mathbf{W} is the $N \times N$ contiguity matrix;

\mathbf{X} is the $N \times q$ matrix of observations of the explanatory variables;

$\tilde{\boldsymbol{\epsilon}}$ is the $N \times 1$ vector of stochastic disturbances;

ρ is the spatial dependence parameter measuring the average influence of contiguous observations on y ;

$\boldsymbol{\gamma}$ is the $p \times 1$ vector of regression coefficients of the explanatory variables.

We adopt the SEM convention to focus on variables, ignoring the units of observation. That is, y is taken to represent the dependent variable as such and is written as a scalar (1×1) rather than $N \times 1$ vector $\tilde{\mathbf{y}}$. This leads to

$$y = \rho y_w + \boldsymbol{\gamma}' \mathbf{x} + \epsilon, \quad (15)$$

which contains the spatially lagged dependent variable y_w (for the time being we ignore the transformation $\mathbf{W}\tilde{\mathbf{y}}$).

We now turn to the formulation of (15) in SEM terms (1). Moreover, we shall apply the SEM model to the well-known example, relating crime to income and housing for 49 contiguous neighborhoods in Columbus, Ohio. Its data matrix \mathbf{Y} and contiguity matrix \mathbf{W} were obtained from Anselin (1988) and from website <http://www.spatial-econometrics.com>. The spatial SEM analyses are performed by the ML option of SEM programs Mx and LISREL 8.

The 49×5 data matrix consists of the five columns y (crime), y_w (spatially weighted crime) x_1 (income), x_2 (housing value), and 1 (unit variable). Because of the presence of the unit variable, the sample moment matrix \mathbf{S} has the variable means in the last row and last column. In a model with observables only, $\mathbf{\Lambda} = \mathbf{I}$ and $\mathbf{\Theta} = \mathbf{0}$ and model implied moment matrix $\mathbf{\Sigma} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{\Psi}(\mathbf{I} - \mathbf{B}')^{-1}$ involves only \mathbf{B} and $\mathbf{\Psi}$ which read:

$$\mathbf{B} = \begin{array}{ccccc} & y & y_w & x_1 & x_2 & 1 \\ \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right. & \begin{array}{c} \rho \\ 0 \\ 0 \\ 0 \\ 0 \end{array} & \begin{array}{c} \gamma_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} & \begin{array}{c} \gamma_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} & \begin{array}{c} \gamma_0 \\ \mu_{y_w} \\ \mu_{x_1} \\ \mu_{x_2} \\ 0 \end{array} & \begin{array}{c} y \\ y_w \\ x_1 \\ x_2 \\ 1 \end{array} \end{array}$$

$$\Psi = \begin{bmatrix} y & y_W & x_1 & x_2 & 1 \\ \sigma^2 & & & & \\ 0 & \sigma_{y_W}^2 & & & \\ 0 & \sigma_{y_W, x_1} & \sigma_{x_1}^2 & & \\ 0 & \sigma_{y_W, x_2} & \sigma_{x_2, x_1} & \sigma_{x_2}^2 & \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} y \\ y_W \\ x_1 \\ x_2 \\ 1 \end{matrix}$$

Moreover, in a model with observables only, the estimated covariance matrix of the explanatory variables (y_W , x_1 , and x_2) is equal to the corresponding sample covariance matrix. The total number of parameters in matrices \mathbf{B} and Ψ (including the coefficient of the unit variable 1) is 15, while also the number of nonidentical elements in the 5×5 sample moment matrix \mathbf{S} is 15. Hence, the model is just identified.

Note that estimating the model by means of the ML option of the SEM program, without special measures taken, results in an estimator, called Lag-OLS by Anselin (1988), which is biased and inconsistent. The reason is that the SEM loglikelihood function (6) as well as the adapted forms (8) and (9) are incomplete. They are based on a transformation of standard normal variates into observed variables which does not take into account that the spatial lag variable y_w is in fact a transformation $\mathbf{W}\tilde{\mathbf{y}}$ of another variable y in the model. Specifically, the transformation implicitly used in SEM is

$$\mathbf{\Omega}^{-\frac{1}{2}}(\tilde{\mathbf{y}} - \mathbf{X}\boldsymbol{\gamma}) = \tilde{\boldsymbol{\nu}}, \quad (16)$$

where $\mathbf{\Omega}$ is a diagonal $N \times N$ matrix with the variances σ^2 on the diagonal

and $\tilde{\boldsymbol{\nu}}$ a $N \times 1$ vector of standard normal variates, whereas actually the following transformation is applied

$$\boldsymbol{\Omega}^{-\frac{1}{2}}(\mathbf{A}\tilde{\boldsymbol{y}} - \mathbf{X}\boldsymbol{\gamma}) = \tilde{\boldsymbol{\nu}} \quad \text{with} \quad \mathbf{A} = \mathbf{I} - \rho\mathbf{W} . \quad (17)$$

This leads via the Jacobian of the transformation to the introduction of an extra additive component $\ln|\mathbf{A}|$ into the loglikelihood (9), which takes in this univariate case the form

$$\begin{aligned} \ell(\boldsymbol{\theta}|\tilde{\boldsymbol{y}}) &= \ln |\boldsymbol{\Omega}^{-\frac{1}{2}}\mathbf{A}| - \frac{1}{2}(\mathbf{A}\tilde{\boldsymbol{y}} - \mathbf{X}\boldsymbol{\gamma})'\boldsymbol{\Omega}^{-1}(\mathbf{A}\tilde{\boldsymbol{y}} - \mathbf{X}\boldsymbol{\gamma}) - \frac{N}{2} \ln 2\pi \\ &= \ln |\mathbf{A}| - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - \mu_i)^2 - \frac{N}{2} \ln 2\pi . \end{aligned} \quad (18)$$

(18) clearly shows that the component $\ln|\mathbf{A}|$ is added to the standard univariate loglikelihood³. Fortunately, the flexibility of the SEM program Mx with its matrix algebraic toolbox and user defined fit function option allows this component to be defined and added to the standard fit function. It means that the spatial lag model can be directly estimated in one run of the program.

In Table 1 the results are compared with those given by Anselin (1988) and Anselin and Bera (1999)⁴. The differences between the values obtained

³As it minimizes $\chi^2 = (N-1)FM = -2(\frac{N-1}{N})[\ell(\boldsymbol{\theta}|\mathbf{Y}) + \text{constant}]$ the correction to be applied to obtain the maximum likelihood solution by means of the Mx program is

$$-2(\frac{N-1}{N}) \ln |\mathbf{A}| . \quad (19)$$

⁴In addition to the parameter estimates the Mx program also computes likelihood based confidence intervals for the parameters (Neale & Miller, 1997). However, we restrict ourselves to point estimates and ignore the confidence intervals.

from the SEM-Mx procedure and Anselin's two-stage procedure are very small and within rounding errors.

	Spatial Lag		Spatial Error	
	SEM-Mx	Anselin	SEM-Mx	Anselin
ρ	0.4314	0.431		
λ			0.5622	0.562
γ_1	-1.0307	-1.032	-0.9403	-0.941
γ_2	-0.2660	-0.266	-0.3022	-0.302
γ_0	45.0568	45.079	59.8791	59.893
σ^2	95.5037	95.495	95.5683	95.575
$\ell(\boldsymbol{\theta} \mathbf{Y})$	-165.4127	-165.408	-166.4006	-166.398
$-2\left(\frac{N-1}{N}\right) \ln \mathbf{A} $	2.2871		4.1899	
corrected χ^2	3.0358		8.0796	
df	0		2	

Table 1: ML estimates of the spatial lag and spatial error models for the Columbus, Ohio, crime data set by the direct SEM-Mx procedure and the two-stage and iterative procedures of Anselin (1988)

Spatial error model in SEM

Although the spatial error model is more complicated because of the implied nonlinearities, its estimation by the flexible nonlinear SEM program Mx is technically not more difficult than for the spatial lag model. We shall illustrate the procedure again by means of the Columbus, Ohio,

crime data. The spatial error model reads:

$$\tilde{\mathbf{y}} = \mathbf{X}\boldsymbol{\gamma} + \tilde{\boldsymbol{\epsilon}} \quad \text{with} \quad \tilde{\boldsymbol{\epsilon}} = \lambda\mathbf{W}\tilde{\boldsymbol{\epsilon}} + \tilde{\boldsymbol{\zeta}}, \quad (20)$$

or:

$$\tilde{\mathbf{y}} = \lambda\mathbf{W}\tilde{\mathbf{y}} + \mathbf{X}\boldsymbol{\gamma} - \lambda\mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \tilde{\boldsymbol{\zeta}}, \quad (21)$$

and in variable formulation:

$$y = \lambda y_w + \boldsymbol{\gamma}'\mathbf{x} - \lambda\boldsymbol{\gamma}'\mathbf{x}_w + \zeta. \quad (22)$$

For the Columbus, Ohio, crime data set (22) becomes

$$y = \lambda y_w + \gamma_1 x_1 + \gamma_2 x_2 - \lambda\gamma_1 x_{w,1} - \lambda\gamma_2 x_{w,2} + (1 - \lambda)\gamma_0 + \zeta. \quad (23)$$

So, the model contains 7 variables, including the unit variable. The SEM

matrices \mathbf{B} and $\boldsymbol{\Psi}$ become

$$\mathbf{B} = \begin{bmatrix} & y & y_w & x_1 & x_2 & x_{w,1} & x_{w,2} & 1 \\ \left[\begin{array}{ccccccc} 0 & \lambda & \gamma_1 & \gamma_2 & -\lambda\gamma_1 & -\lambda\gamma_2 & (1-\lambda)\gamma_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu_{y_w} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu_{x_1} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu_{x_2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu_{x_{w,1}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu_{x_{w,2}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] & \begin{array}{l} y \\ y_w \\ x_1 \\ x_2 \\ x_{w,1} \\ x_{w,2} \\ 1 \end{array} \end{bmatrix},$$

$$\Psi = \begin{bmatrix} y & y_W & x_1 & x_2 & x_{W,1} & x_{W,2} & 1 \\ \sigma^2 & & & & & & \\ 0 & \sigma_{y_W}^2 & & & & & \\ 0 & \sigma_{x_1, y_W} & \sigma_{x_1}^2 & & & & \\ 0 & \sigma_{x_2, y_W} & \sigma_{x_2, x_1} & \sigma_{x_2}^2 & & & \\ 0 & \sigma_{x_{W,1}, y_W} & \sigma_{x_{W,1}, x_1} & \sigma_{x_{W,1}, x_2} & \sigma_{x_{W,1}}^2 & & \\ 0 & \sigma_{x_{W,2}, y_W} & \sigma_{x_{W,2}, x_1} & \sigma_{x_{W,2}, x_2} & \sigma_{x_{W,2}, x_{W,1}} & \sigma_{x_{W,2}}^2 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} y \\ y_W \\ x_1 \\ x_2 \\ x_{W,1} \\ x_{W,2} \\ 1 \end{matrix}.$$

The matrix \mathbf{B} has the following entries to be estimated: $\lambda, \gamma_1, \gamma_2, -\lambda\gamma_1, -\lambda\gamma_2, (1 - \lambda)\gamma_0, \mu_{y_W}, \mu_{x_1}, \mu_{x_2}, \mu_{x_{W,1}}$ and $\mu_{x_{W,2}}$. Note that there are three nonlinearly restricted entries $-\lambda\gamma_1, -\lambda\gamma_2, (1 - \lambda)\gamma_0$ ⁵.

Disregarding all (co)variances to be estimated with respect to the independent variables that are immediately given by the sample covariance matrix, there are 5 parameters left ($\lambda, \gamma_1, \gamma_2, \gamma_0, \sigma^2$) to be estimated, just as many as for the spatial lag model. However, the sample moment matrix contains 7 nonidentical elements to estimate them. So, if all parameters are identified individually, the model as a whole is overidentified with $df = 2$.⁶

The transformation the loglikelihood should be based on, is easily

⁵The entry for the constant $(1 - \lambda)\gamma_0$ (coefficient of the unit variable 1) is explained by the fact that spatially weighting a constant variable 1 gives 1 again, so that γ_0 (for $x_0 = 1$) and the spatial correction $-\lambda\gamma_0$ (for $x_{W,0} = 1$) are estimated in combination for the single unit variable 1.

⁶The reason for overidentification is that the SEM approach introduces three new transformed variables, while under the spatial lag model only one new transformed variable is introduced.

derived from (21):

$$\mathbf{\Omega}^{-\frac{1}{2}}\mathbf{A}(\tilde{\mathbf{y}} - \mathbf{X}\boldsymbol{\gamma}) = \tilde{\boldsymbol{\nu}} \quad \text{with} \quad \mathbf{A} = \mathbf{I} - \lambda\mathbf{W} . \quad (24)$$

Although it is different from (17), it has the same Jacobian as the spatial lag model, when the lag parameter ρ is replaced by the error parameter λ . It means that to obtain the ML solution, the correction in the minimization of χ^2 should take place analogously to matrix (19). Table 1 shows that, again, the values obtained by the direct SEM-Mx procedure are close to the estimates obtained by Anselin.

While we have shown in this section that the standard spatial lag and error models can directly be estimated in one run by means of the SEM program Mx, other SEM programs like LISREL 8 can also be used to do the job in an iterative stepwise fashion as follows. First get a starting value for the spatial parameter ρ or λ by estimating all model parameters, including the spatial parameter ρ or λ , by means of OLS. Next fix the spatial parameter at the OLS-value found and compute the χ^2 - correction (19) by any matrix algebraic program as, for example, GAUSS. Increase or decrease the spatial parameter value stepwise until the minimum corrected χ^2 is found.

4. Spatial lag model for a latent dependent variable

The model in this section will be coined latent spatial lag model. It

differs from the standard spatial lag model (see section 3) in that the regressand is a latent variable (MIMIC model), measured by several indicators. Spatial dependence applies to the latent variable instead of the observed variable. So, the structural equation looks like (15) but has observed y replaced by latent η :

$$\eta = \rho\eta_w + \boldsymbol{\gamma}'\mathbf{x} + \zeta . \quad (25)$$

It is completed by measurement equations

$$\mathbf{y} = \boldsymbol{\Lambda}\eta + \boldsymbol{\epsilon} \quad \text{and} \quad \mathbf{y}_w = \boldsymbol{\Lambda}\eta_w + \boldsymbol{\epsilon}_w \quad (26)$$

with m observed variables in \mathbf{y} as well as m observed variables in \mathbf{y}_w . We assume $\boldsymbol{\Lambda}$ to be equal for the original \mathbf{y} and the lagged \mathbf{y}_w , which seems desirable to achieve measurement invariance for latent η and lagged latent η_w . However, if wanted, this assumption is easily relaxed.

In observation unit form the model becomes

$$\tilde{\boldsymbol{\eta}} = \rho\mathbf{W}\tilde{\boldsymbol{\eta}} + \mathbf{X}\boldsymbol{\gamma} + \tilde{\boldsymbol{\zeta}} , \quad (27)$$

$$\tilde{\mathbf{y}} = \tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\eta}} + \tilde{\boldsymbol{\epsilon}} \quad \text{and} \quad \tilde{\mathbf{y}}_w = \tilde{\boldsymbol{\Lambda}}\mathbf{W}\tilde{\boldsymbol{\eta}} + \tilde{\boldsymbol{\epsilon}}_w , \quad (28)$$

where $\tilde{\boldsymbol{\eta}}$ and $\tilde{\boldsymbol{\zeta}}$ are $N \times 1$ but $\tilde{\mathbf{y}}, \tilde{\boldsymbol{\epsilon}}, \tilde{\mathbf{y}}_w$ and $\tilde{\boldsymbol{\epsilon}}_w$ are $Nm \times 1$ with

$$\tilde{\boldsymbol{\Lambda}}_{Nm \times N} = \mathbf{I}_{N \times N} \otimes \boldsymbol{\Lambda}_{m \times 1} ,$$

$$\tilde{\mathbf{y}}_w = \tilde{\mathbf{W}}\tilde{\mathbf{y}} \quad \text{for} \quad \tilde{\mathbf{W}}_{Nm \times Nm} = \mathbf{W}_{N \times N} \otimes \mathbf{I}_{m \times m} .$$

It should be noted that we assumed One first derives

$$\begin{aligned}\tilde{\mathbf{y}} &= \rho\tilde{\Lambda}\mathbf{W}\tilde{\boldsymbol{\eta}} + \tilde{\Lambda}\mathbf{X}\boldsymbol{\gamma} + \tilde{\Lambda}\tilde{\boldsymbol{\zeta}} + \tilde{\boldsymbol{\epsilon}} \\ &= \rho\tilde{\mathbf{y}}_w - \rho\tilde{\boldsymbol{\epsilon}}_w + \tilde{\Lambda}\mathbf{X}\boldsymbol{\gamma} + \tilde{\Lambda}\tilde{\boldsymbol{\zeta}} + \tilde{\boldsymbol{\epsilon}}\end{aligned}$$

and then

$$(\mathbf{I} - \rho\tilde{\mathbf{W}})\tilde{\mathbf{y}} = \tilde{\Lambda}\mathbf{X}\boldsymbol{\gamma} + \tilde{\Lambda}\tilde{\boldsymbol{\zeta}} + \tilde{\boldsymbol{\epsilon}} - \rho\tilde{\boldsymbol{\epsilon}}_w \quad ,$$

which for $\tilde{\mathbf{A}} = \mathbf{I} - \rho\tilde{\mathbf{W}}$ leads to the transformation from m -dimensional $\tilde{\boldsymbol{\nu}}$ to m -dimensional $\tilde{\mathbf{y}}$

$$\begin{aligned}\boldsymbol{\Omega}^{-\frac{1}{2}}(\tilde{\mathbf{A}}\tilde{\mathbf{y}} - \tilde{\Lambda}\mathbf{X}\boldsymbol{\gamma}) &= \tilde{\boldsymbol{\nu}} \quad \text{for} \\ \boldsymbol{\Omega} &= \mathbf{I} \otimes (\boldsymbol{\Lambda}\sigma_{\zeta_1}^2\boldsymbol{\Lambda}' + \boldsymbol{\Theta}_\epsilon + \rho^2\boldsymbol{\Theta}_{\epsilon_w} - \rho\boldsymbol{\Theta}_{\epsilon,\epsilon_w} - \rho\boldsymbol{\Theta}_{\epsilon_w,\epsilon}) .\end{aligned}\quad (29)$$

Via the Jacobian of the transformation

$$\mathbf{J} = |\boldsymbol{\Omega}^{-\frac{1}{2}}| |\tilde{\mathbf{A}}|$$

it is seen that an extra component $\ln |\tilde{\mathbf{A}}| = m \ln |\mathbf{A}|$ is to be added to the loglikelihood, which is m (the number of indicators of the latent variable) times the correction in the standard spatial lag model, and that in SEM covariance matrix $\boldsymbol{\Sigma}$ the standard form $\boldsymbol{\Lambda}\sigma_{\zeta_1}^2\boldsymbol{\Lambda}' + \boldsymbol{\Theta}_\epsilon$ should be replaced by the corrected form between parentheses in $\boldsymbol{\Omega}$ in (29).

Instead of estimating the combination $\boldsymbol{\Theta}_\epsilon^*$

$$\boldsymbol{\Theta}_\epsilon^* = \boldsymbol{\Theta}_\epsilon + \rho^2\boldsymbol{\Theta}_{\epsilon_w} - \rho\boldsymbol{\Theta}_{\epsilon,\epsilon_w} - \rho\boldsymbol{\Theta}_{\epsilon_w,\epsilon}\quad (30)$$

as in standard SEM, one should estimate Θ_ϵ in conjunction with the other components in (30) in terms of parameter ρ (to keep ρ within appropriate bounds, one should then also specify $0 < \rho < 1$). However, because, in general, (30) does not put extra constraints on ρ , first estimating Θ_ϵ^* and then afterwards computing its components in (30), in general, will give the same results. It should be noted also, that the correlation between measurement errors $\tilde{\epsilon}_w$ and $\tilde{\epsilon}$, coming from different observation units, will be nonexistent or very small and that the measurement variances in Θ_{ϵ_w} will be much smaller than in Θ_ϵ . The reason for the latter is that the relative measurement error of linear combinations of indicators is much smaller than of the single indicators (Lord & Novick, 1968, pp. 85-87). The consequence is that in practice the difference between Θ_ϵ^* and Θ_ϵ will be small, while also knowledge of Θ_ϵ^* (measurement error variance under spatial dependence) will be more relevant than of the rather theoretical Θ_ϵ (measurement error variance under the assumption of no spatial dependence).

We conclude that for the latent spatial lag model the correction to the χ^2 is

$$-2\left(\frac{N-1}{N}\right)m \ln |\mathbf{A}| \quad (31)$$

and $\Lambda\sigma_{\zeta_1}^2 \Lambda' + \Theta_\epsilon$ in Σ is to be corrected by adding

$$\rho^2 \Theta_{\epsilon_w} - \rho \Theta_{\epsilon, \epsilon_w} - \rho \Theta_{\epsilon_w, \epsilon} . \quad (32)$$

We applied the latent spatial lag model to the Columbus, Ohio, crime data, where we considered Income and Housing as indicators of a latent variable Social Economic Situation and analyzed the effect of Crime on Social Economic Situation. The estimation results of the SEM-Mx procedure are given in Table 2. We show the estimated SEM matrices \mathbf{B} , $\mathbf{\Psi}$, $\mathbf{\Lambda}$, and $\mathbf{\Theta}^*$. The model χ^2 was 9.277 (noncorrected value 8.639 + correction 0.638) with degrees of freedom $df = 5$ ($p = 0.099$), which indicates that the model is fitting very well.

$$\begin{aligned}
\hat{\mathbf{B}} &= \begin{bmatrix} & \text{SES} & \text{Lagged SES} & \text{Crime} & 1 \\ \begin{bmatrix} - & 0.170 & -0.228 & 19.9 \\ - & - & - & 35.1 \\ - & - & - & 14.7 \\ - & - & - & - \end{bmatrix} & \text{SES} \\ & \text{Lagged SES} \\ & \text{Crime} \\ & 1 \end{bmatrix} \\
\hat{\mathbf{\Psi}} &= \begin{bmatrix} 5.043 & & & & \\ - & 16.48 & & & \\ - & -39.54 & 274.2 & & \\ - & - & - & 1 & \end{bmatrix} \begin{matrix} \text{SES} \\ \text{Lagged SES} \\ \text{Crime} \\ 1 \end{matrix} \\
\hat{\mathbf{\Lambda}} &= \begin{bmatrix} 1 & - & - & - \\ 1.85 & - & - & 19.7 \\ - & 1 & - & - \\ - & 1.85 & - & 19.7 \\ - & - & 1 & - \\ - & - & - & 1 \end{bmatrix} \begin{matrix} \text{Income} \\ \text{Housing} \\ \text{Lagged Income} \\ \text{Lagged Housing} \\ \text{Crime} \\ 1 \end{matrix}
\end{aligned}$$

	Income	Housing	Lagged Income	Lagged Housing	Crime	1	
$\hat{\Theta}^*$ =	10.58	—	—	—	—	—	Income
	—	216.3	—	—	—	—	Housing
	—	—	0.86	—	—	—	Lagged Income
	—	—	—	62.8	—	—	Lagged Housing
	—	—	—	—	—	—	Crime
	—	—	—	—	—	—	1

Table 2: Estimated SEM matrices of the latent spatial lag model applied to the Columbus, Ohio, crime data set (Anselin, 1988)

Both the rather low latent spatial lag parameter (0.170) as well as the quite substantial negative effect exerted by Crime on the SES of the neighborhood (-0.288), were tested for significance by the χ^2 -difference test. The spatial lag parameter was found to be nonsignificant ($\chi^2_{df} = 1.528$ with $df = 1$, $p = 0.216$) but the effect of Crime was highly significant ($\chi^2_{df} = 27.588$ with $df = 1$, $p < 0.00001$).

The measurement unit (loading) and measurement origin in measuring SES were both estimated to be greater for Housing (1.85 and 19.7, respectively) than for Income. The corrected measurement error variance θ_ϵ^* was estimated to be 10.58 for Income and 216.3 for Housing, corresponding to reliabilities R^2 of 0.667 and 0.352, respectively. This shows that Income is a much more reliable indicator of SES than Housing. The estimates of the uncorrected measurement error variances θ_ϵ of 10.55 and 214.5 were indeed hardly different from the corrected ones and, as expected, the estimates of the lagged measurement error variances θ_{ϵ_w} (0.86 and 62.8) were

much lower than either the corrected or uncorrected ones.

5. Latent variable representation of the spatial lag model

In this section we propose the following SEM for the spatial lag model

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\mathbf{x} + \boldsymbol{\zeta} , \quad (33)$$

$$\mathbf{y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\epsilon} , \quad (34)$$

where

$$\mathbf{B} = \begin{bmatrix} 0 & \beta_{12} \\ \beta_{12} & 0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_0 \\ 0 & 0 & \dots & \mu_{\eta_2} \end{bmatrix}, \quad \boldsymbol{\Psi} = \begin{bmatrix} \sigma_{\zeta_1}^2 & 0 \\ 0 & \sigma_{\zeta_2}^2 \end{bmatrix},$$

$$\mathbf{y} = \begin{bmatrix} y \\ y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}, \quad \boldsymbol{\Lambda}_y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & \lambda_2 \\ \vdots & \vdots \\ 0 & \lambda_k \end{bmatrix}, \quad \boldsymbol{\Theta}_\epsilon = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\epsilon_2}^2 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\epsilon_k}^2 \end{bmatrix},$$

The following observations apply. First we again adopted the SEM convention to specify the model in terms of variables rather than in units of observation. That is y, y_1, y_2, \dots, y_k in \mathbf{y} are written as scalars (1×1) rather than $N \times 1$ vectors. Secondly, the lagged observed variables are included in the vector \mathbf{y} and are taken as indicators of the latent variable η_2 . We assume that the spatially lagged observed variables are selected on the basis of theoretical or ad hoc considerations. Selection is done by means of

selection or weight matrices \mathbf{W}_i as follows:

$$\begin{aligned}\tilde{\mathbf{y}}_1 &= \mathbf{W}_1 \tilde{\mathbf{y}} \\ \tilde{\mathbf{y}}_2 &= \mathbf{W}_2 \tilde{\mathbf{y}} \\ &\vdots \\ \tilde{\mathbf{y}}_k &= \mathbf{W}_k \tilde{\mathbf{y}}\end{aligned}\tag{35}$$

That is, \mathbf{W}_1 selects the values for the first indicator y_1 , \mathbf{W}_2 for the second indicator y_2 , etc. Third, simultaneity bias due to the interaction between the dependent variable and its lags is controlled for by the matrix \mathbf{B} . Specifically, β_{12} represents the impacts of the latent spill-over variable on the dependent variable (i.e. β_{12} corresponds to ρ in the standard lag model), whereas β_{21} represents the reverse effect. Fourth, it is assumed again that the unit variable is included as the last variable in \mathbf{x} , having the regression intercept γ_0 (for $\eta_1 = y$) as associated coefficient and specifying also the mean μ_{η_2} of the spill-over variable.

Below we illustrate the proposed approach using the Columbus crime data set again. Observe that the selection matrices (35) are based on distance. That is, for each region the three nearest neighbours are selected. We defined \mathbf{W}_1 as the selector of the nearest contiguous neighbor, \mathbf{W}_2 of the next nearest contiguous neighbor and \mathbf{W}_3 of the third nearest contiguous neighbor. Seven units had only two contiguous neighbors and for these cases we let \mathbf{W}_3 select the nearest non-contiguous neighbor.

The estimation results of the SEM-Mx procedure are given in Table 3. We show the estimated SEM matrices \mathbf{B} , $\mathbf{\Psi}$, and $\mathbf{\Lambda}$ (formulation (1)-

(2)). The measurement error variances of Crime-neighbor2 and Crime-neighbor3 were estimated to be 185.4 and 202.9 (corresponding to R^2 's of 0.881 and 0.873, respectively). The model χ^2 was 27.8 with degrees of freedom $df = 10$ ($p = 0.004$). We observe that the overall fit as indicated by χ^2 is reasonable to good with a χ^2/df of less than 3 and a CFI (comparative Fit Index) of 0.977. This conclusion is supported by the modification indices of parameters eligible to be freed, which are all rather small.

The measurement model ($\hat{\Lambda}$) shows that the nearest neighbor is the most important indicator for the latent variable spill-over. Moreover, the impacts of the second and third neighbors are virtually equal. (This could be formally tested by imposing an equality constraint for the corresponding coefficients, re-estimating the model and calculating the difference in χ^2 values.)

$$\hat{B} = \begin{bmatrix} \text{Crime} & \text{Crime-neighbor} & \text{Income} & \text{Housing} & 1 \\ - & 0.073 & -1.489 & -0.279 & 64.372 \\ - & (0.745) & (-4.393) & (-2.798) & (6.483) \\ 0.548 & - & - & - & 20.420 \\ (3.622) & - & - & - & (3.625) \\ - & - & - & - & 14.377 \\ - & - & - & - & (17.643) \\ - & - & - & - & 38.436 \\ - & - & - & - & (14.571) \\ - & - & - & - & - \end{bmatrix} \begin{matrix} \text{Crime} \\ \text{Crime-neighbor} \\ \text{Income} \\ \text{Housing} \\ 1 \end{matrix}$$

$$\hat{\Psi} = \begin{bmatrix} 112.4 \\ (30.2) \\ - & 166.4 \\ - & (34.2) \\ - & - & 31.9 \\ - & - & (4.9) \\ - & - & 51.6 & 334.0 \\ - & - & (3.1) & (4.9) \\ - & - & - & - & 1 \end{bmatrix} \begin{matrix} \text{Crime} \\ \text{Crime-neighbor} \\ \text{Income} \\ \text{Housing} \\ 1 \end{matrix}$$

$$\hat{\Lambda} = \begin{bmatrix} 1 & - & - & - & - \\ - & 1 & - & - & - \\ - & 0.863 \\ - & (18.8) & - & - & - \\ - & 0.871 \\ - & (18.2) & - & - & - \\ - & - & 1 & - & - \\ - & - & - & 1 & - \\ - & - & - & - & 1 \end{bmatrix} \begin{matrix} \text{Crime} \\ \text{Crime-neighbor1} \\ \text{Crime-neighbor2} \\ \text{Crime-neighbor3} \\ \text{Income} \\ \text{Housing} \\ 1 \end{matrix}$$

Table 3: Estimated SEM matrices of the lag model applied to the Columbus, Ohio, crime data set (Anselin, 1988); t -values between parentheses

From the structural model it follows that the overall spatial spill-over effect is positive, though insignificant. Particularly, it is small compared to the own housing effect and very small relative to the own income effect. Although they are not quite comparable due to amongst others differences in the way spatial spill-over is measured (contiguity-based in Anselin; distance based in the present case), the coefficients for housing are very close to those presented in Table 1. For income the coefficients in Table 1 exceed those obtained here. Finally we observe that the coefficient in the second structural equation representing the impact of the dependent variable on

the spill-over variable is large compared to the reverse effect. This is most likely due to the absence of other explanatory variables in this equation. Further research on this issue is needed.

6. Conclusion

In this paper we have considered structural equations (SEM) approaches to spatial dependence models. As a first step, we adapted the standard SEM likelihood function such that the standard lag and error models can be estimated in a straightforward fashion by the standard SEM software packages. Application to the Columbus, Ohio, crime data set (Anselin, 1998) showed that by these packages virtually the same estimates are obtained as by the standard software for spatial dependence models. The SEM software package Mx has the main advantage that it allows one-time estimation rather than the iterative and two-stage procedures used in standard spatial modelling software. Moreover, SEM allows handling of nonlinearities in a straightforward fashion.

In the second part of the paper we used the SEM framework to model the spatial configuration for the spatial lag model. First, we generalized the spatial lag model to latent dependent variables, of which the observed lags are indicators. Next, the conventional way of representing the spatial model by means of a weight matrix W describing the spatial arrangement

of the units of observation was replaced by latent variables representing spatial dependence of which the observed values of the neighbouring spatial units are indicators. This approach is both more flexible and more informative than modelling by means of an a priori given spatial weight matrix W . In addition we presented a Full Information Maximum Likelihood Estimator to control for endogeneity bias due to the interaction between the dependent variable and its lags. The models were applied to the Columbus, Ohio, crime data set. Although for several parameter estimates were obtained in line with the results obtained by Anselin (1998), further research, particularly Monte Carlo simulation, is needed.

The SEM approach to spatial dependence models has several potential advantages which are worthwhile further exploring. Particularly, the SEM approach allows straightforward application to systems of equations. Secondly, it allows the introduction of several dynamic SEM features into spatial modelling.

Literature

Anselin, L., (1988). *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic Publishers.

Anselin, L., & Bera, A.K. (1999). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah &

- D.E.A. Giles (Eds.), *Handbook of applied economic statistics* (pp. 237-289). New York: Marcel Dekker.
- Anselin, L., & Hudak, A.K. (1992). Spatial econometrics in practice: A review of software options. *Regional Science and Urban Economics*, 22, 509-536.
- Blalock, H.M. (Ed.) (1971). *Causal models in the social sciences*. London: MacMillan.
- Carnap, R. (1936). Testability and meaning. *Philosophy of science*, 3, 419-471.
- Folmer, H., (1986). *Regional economic policy*. Dordrecht: Kluwer Academic Publishers.
- Hempel, C.G. (1958). The theoretician's dilemma. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science*, vol. 2 (pp. 37-98). Minneapolis: University of Minnesota Press.
- Jöreskog, K.G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software International.
- Lord, F.M., & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley .
- Neale, M.C., Boker, S.M., Xie, G., & Maes, H.H. (2003). *Mx: Statistical Modeling* (6th ed.). Richmond, VA: Department of Psychiatry.
- Neale, M.C., & Miller, M.B. (1997). The use of likelihood-based confidence

intervals in genetic models. *Behavior Genetics*, 27,113-120.