SPECIFICATION AND ESTIMATION OF SPATIAL LINEAR REGRESSION MODELS:
Monte Carlo Evaluation of Pre-Test Estimators

Raymond Florax, Henk Folmer*

forthcoming in *Regional Science and Urban Economics*

* Part of this research was accomplished while the first author was a research associate at the Center for Higher Education Policy Studies (CHEPS), University of Twente, The Netherlands. The authors are now both associated with Wageningen Agricultural University, Department of General Economics, P.O. Box 8130, 6700 EW Wageningen, The Netherlands.
A preliminary version of this paper was presented at the workshop on Statistical Modeling of the Commission on Mathematical Models of the International Geographical Union, Boston, Massachusetts, November 8, 1990, and at the North American Meetings of the Regional Science Association, New Orleans, November 8, 1991.
The authors would like to acknowledge the assistance of Marc Loman in writing the computer program, and the useful suggestions of Luc Anselin, Pierre van Mouche and two anonymous referees.

SPECIFICATION AND ESTIMATION OF SPATIAL LINEAR REGRESSION MODELS:
Monte Carlo Evaluation of Pre-Test Estimators

Abstract

Spatially correlated residuals lead to various serious problems in applied spatial research. In this paper several conventional specification and estimation procedures for models with spatially dependent residuals are compared with alternative procedures. The essence of the latter is a search procedure for spatially lagged variables. By incorporating the omitted spatially lagged variables into the model spatially dependent residuals may be remedied, in particular if the spatial dependence is substantive. The efficacy of the conventional and alternative procedures in small samples will be investigated by means of Monte Carlo techniques for an irregular lattice structure.

1. Introduction

Spatial dependence among the disturbances of spatial models is a serious problem in empirical research. In particular, the frequently applied ordinary least squares (OLS) estimator is inefficient, the estimator of the residual variance is biased, the values of the estimated R^2 are inflated and inference procedures are invalid [cf. Cliff and Ord (1981)]. Furthermore, spatially correlated residuals affect the properties of tests regarding model selection and heteroscedasticity [Anselin and Griffith (1988), Anselin (1990)]. Conversely, heteroscedasticity affects tests on spatial dependence [cf. Anselin and Rey (1991)].

As the drawbacks of spatially correlated disturbances are severe, it is not surprising that various attempts have been made to handle this problem. In this connection the following procedures can be distinguished:

- remedial action which consists of some kind of transformation of the sample observations, leading to estimated generalized least squares (EGLS), such as the Cochrane-Orcutt [e.g. Hordijk (1974)] or the Durbin estimator [e.g. Ord (1975), Anselin (1988)], or to variables with the autoregressive components filtered out [Getis (1990)];
- maximum likelihood estimation, which has been applied to various spatial dependence models, such as spatial AR and MA models [e.g. Cliff and Ord (1981), Anselin (1988)]; and
- adjustments in the context of model specification, either through spatial expansion [Casetti (1972), Jones (1983), Casetti and Jones (1988)], or spatial adaptive filtering [Foster and Gorr (1983, 1984)].

Various test statistics for spatial correlation among the residuals of linear regression models have been developed, such as the Moran coefficient, Geary's coefficient, the Cliff and Ord statistic and Lagrange Multiplier tests [cf. e.g. Cliff and Ord (1972, 1973, 1981), Hordijk (1974), Anselin (1988), Getis (1991)]. The Moran coefficient has been found to be easily applicable and to perform reasonably well in a large variety of situations. Yet, recent results based on econometric simulation studies [Anselin and Rey (1991), Florax and Folmer (1991)] cast some doubt about whether the Moran coefficient is suitable for the detection of substantive spatial dependence, i.e. spatial correlation among the residuals as a result of erroneously omitted, spatially lagged, explanatory variables.

Florax and Folmer (1991) show that, for an irregular lattice structure, the Moran coefficient has power against spatial autoregressive errors as well as in the case of models with an erroneously omitted, spatially lagged, dependent variable. With the spatially lagged, exogenous variables erroneously omitted, however, the use of the Moran coefficient may easily lead to excessive Type-II errors, depending on the variance of the error term. These conclusions are demonstrated to hold for inference based on the common asymptotic properties of the Moran coefficient as well as for non-parametric approaches, such as the bootstrap and the *permutation procedure* [see Cliff and Ord (1981), Folmer and Fischer (1984) and Folmer (1986) for these non-parametric resampling techniques in spatial analysis]. The results of Anselin and Rey (1991) among other things show that similar conclusions hold for regular lattice structures (except for omitted, spatially lagged, exogenous variables and resampling techniques which they did not investigate). A strategy to detect substantive spatial dependence based solely on the use of Moran's *IR* will therefore not be uniformly effective.

The Lagrange Multiplier tests for spatially correlated residuals (*LMERR*) and for omitted, spatially lagged, dependent variables (*LMLAG*) have not been tailored to the detection of erroneously omitted, exogenous variables either. Hence, at the moment no adequate test statistics for the detection of erroneously omitted, spatially lagged, exogenous variables are available. As the consequences of this kind of specification error are serious (i.e. the coefficient estimator is biased, the disturbance variance is overestimated, and inference procedures are invalid), it is highly desirable to develop procedures to detect and remedy it.

The purpose of the present paper is to present and analyze specification and estimation procedures designed to identify erroneously omitted, spatially lagged, systematic variables; to include these variables into the model, and to resort to data transformation or ML estimation only when there is sufficient evidence that the spatially correlated residuals are a consequence of spatial correlation among the variables represented by the error term. This approach to handling misspecification of spatial regression models is referred to as spatial variable expansion (*SVE*) in this paper. Two alternatives of the spatial variable expansion procedure are considered. One is based on the Moran coefficient for regression residuals (*IR*) and *F* tests on omitted, spatially lagged, explanatory variables (*FOV*). The only estimator applied in that case is ordinary least squares (OLS). A justification for this procedure is given in section four. The other alternative makes use of the Lagrange Multiplier test statistics for spatially lagged dependent variables (*LMLAG*) and for spatially correlated errors (*LMERR*) in conjunction with *FOV* or Likelihood Ratio (*LR*) tests to identify erroneously omitted, spatially lagged, explanatory variables. A limited variable expansion method, which only leads to expansion by the spatially lagged dependent variable, is also considered.

In detail the organization of this paper is as follows. In section two a taxonomy of spatial dependence models is presented. Section three discusses some of the more traditional solutions to the spatial autocorrelation problem, viz. EGLS and ML estimation. In section four an overview of the variable expansion procedures will be given. In section five the performance of the variable expansion procedure is analyzed by means of a simulation study for an irregular lattice structure. Different forms of misspecification as well as the model with autoregressive disturbances are investigated. The simulations also include the EGLS and ML estimators so as to provide a bench-mark for the variable expansion procedures. In section six some concluding remarks wind up this paper.

2. A taxonomy of linear spatial dependence models

In order to provide a framework for the analysis of various forms of spatial dependence consider the following general linear regression model for spatial cross-sections r (= 1, 2, ..., R):

$$y = \zeta \mathbf{W} y + \mathbf{X} \beta + \mathbf{W} \mathbf{X}^{*} \rho + \varepsilon \qquad (1.1)$$

$$\varepsilon = \lambda \mathbf{W} \varepsilon + \mu \qquad (1.2)$$

where y is the $(R \times 1)$ stochastic dependent variate, $\mathbf{W}$ an a priori specified $(R \times R)$ spatial weights matrix, $\mathbf{X}$ the $(R \times k)$ matrix of non-stochastic regressors, $\mathbf{X}^{*}$ the $(R \times (k\text{-}1))$ matrix of explanatory variables with the constant term deleted, ζ the autocorrelation coefficient, β the $(k \times 1)$ vector of coefficients of the non-weighted independent variables, ρ the $((k\text{-}1) \times 1)$ vector of crosscorrelation coefficients, λ the coefficient of the autoregressive error term, and μ a vector of random errors with $E(\mu) = 0$ and $E(\mu\, \mu') = \sigma^2_{\mu}\, \mathbf{I}$. This general model will be called a mixed regressive-spatial regressive model with autoregressive disturbances. The following observations apply. First, for an unstandardized weights matrix the term $\mathbf{WX}\rho$ instead of $\mathbf{WX}^{*}\rho$ can be included. In that case a regression coefficient is obtained which corresponds to a variable made up of the row sums of the weights matrix.

Secondly, a model with both autoregressive disturbances and substantive spatial dependence would require two (or more) different weights matrices, e.g. $y = \zeta \mathbf{W}_1 y + \mathbf{X}\beta + \mathbf{W}_1\mathbf{X}^*\rho + \varepsilon$ with $\varepsilon = \lambda\mathbf{W}_2\varepsilon + \mu$, because otherwise the $(3 + k + (k\text{-}1))$ unknown parameters $(\zeta,\beta',\rho',\lambda,\sigma^2)'$ are not identified [cf. Anselin (1988)]. For ease of exposition, and because only genuine submodels of the general model will be investigated here, a single weights matrix **W** is used.

At least four special cases can be derived from the model given in (1.1) and (1.2) by putting constraints on the model. These special cases are the following.

- A mixed regressive-spatial autoregressive model ($\rho = (0, ..., 0)'$ and $\lambda = 0$):

$$y = \zeta \mathbf{W} y + \mathbf{X}\beta + \mu \tag{2}$$

  where μ is again a $(R \times 1)$ vector of disturbances with $E(\mu) = 0$ and $E(\mu\,\mu') = \sigma^2_\mu \mathbf{I}$.
- A mixed regressive-spatial crossregressive model ($\zeta = 0$ and $\lambda = 0$):

$$y = \mathbf{X}\beta + \mathbf{W}\mathbf{X}^*\rho + \mu \tag{3}$$

  with μ as given above.
- A mixed regressive-spatial regressive model ($\lambda = 0$) as a result of the combination of models (2) and (3):

$$y = \zeta \mathbf{W} y + \mathbf{X}\beta + \mathbf{W}\mathbf{X}^*\rho + \mu \tag{4}$$

  with μ as before.
- A model with autoregressive disturbances ($\zeta = 0$ and $\rho = (0, ..., 0)'$). Under the assumption that $(\mathbf{I} - \lambda\mathbf{W})$ is invertible and $|\lambda| < 1$ for reasons of stationarity [cf. also Anselin (1982, p. 1025)] this type of model reads as:

$$y = \mathbf{X}\beta + (\mathbf{I} - \lambda\,\mathbf{W})^{-1}\mu \tag{5}$$

  with μ as before.

With respect to the taxonomy of linear spatial dependence models three observations are worth noticing. First, more variety in the taxonomy of spatial models can be obtained by introducing different weights matrices, i.e. weights matrices which differ along with, for instance, the (omitted) variables or with the order of contiguity [cf. Hordijk (1974), Cliff and Ord (1981)]. As a rule, the weights matrix consists of binary contiguity indexes, or is made up of indices based on distance, the relative common perimeter length [Cliff and Ord (1981)], or multidimensional scaling [Gatrell (1979)].

Secondly, the issue of the specification of the weights matrix has received little attention. However, there is evidence in the literature that the *a priori* and exogenous nature of the spatial weights matrix gives rise to two important caveats. First, as has been shown in simulation studies, the power of autocorrelation statistics (e.g. Moran's *IR* and *LM* statistics) is crucially dependent on both the type of weights matrix and whether or not it has been standardized [cf. Stetzer (1982), Anselin and Rey (1991), Florax and Folmer (1991)]. Secondly, misspecification of the weights matrix has an impact on hypothesis testing with respect to spatial dependence among residuals as well as a drawback in terms of bias regarding the estimated coefficients in models with spatially lagged variables. As this issue has been largely disregarded in the literature, it merits further attention [cf. Anselin (1985), Florax and Rey (1992)].

Thirdly, the phenomenon of crossregressive spatial dependence has received only sparse attention in the literature [cf. Folmer (1986, p. 99), for an explicit treatment, and Folmer and Nijkamp (1987) for an application]. The following example shows that it has a real and, depending on the specific context, also a policy relevant meaning. Consider an aggregate regional production function where regional production is treated as a function of *inter alia* the availability of labor. Autocorrelation then implies that regional production in region r is also influenced by regional production in region r', whereas crosscorrelation indicates that regional production in region r is also influenced by the availability of labor in region r' ($r \neq r'$). The influence of both types of spatial causation can of course occur simultaneously. Moreover, there may exist reciprocal influences; that is, from region r to region r' and *vice versa*.

Omission of spatially lagged variables is an important cause for spatially correlated residuals. To illustrate this, assume that the true model is the spatial Durbin or common factor model that includes spatially lagged variables, dependent as well as explanatory [cf. Burridge (1981), Anselin (1990)]. That is:

$$y = \mathbf{X}\beta + \lambda \mathbf{W}(y - \mathbf{X}\beta) + \mu \qquad (6)$$

If the hypothesized model is:

$$y = \mathbf{X}\beta + \varepsilon \qquad (7)$$

then:

$$\varepsilon = \lambda \mathbf{W} (y - \mathbf{X} \beta) + \mu \tag{8}$$

It can be easily seen that the off-diagonal terms of the variance-covariance matrix of ε in general are nonzero. The same holds when, spatially lagged, exogenous variables are erroneously omitted.

It should be observed that spatial autocorrelation among the regression residuals may also imply the presence of non-linear relationships between the dependent variable and one or more explanatory variables, or the presence of spatial correlation among one or more non-systematic variables represented by the error term. Cliff and Ord (1981) give a detailed review of the causes and remedies with regard to these two types of spatial correlation.

In the sequel two types of spatial autocorrelation will be investigated, which may be referred to as substantive spatial dependence and spatial dependence as a nuisance [cf. Doreian (1980), Anselin and Griffith (1988), Anselin and Rey (1990)]. Substantive spatial dependence is meant to imply an autoregressive residual structure due to the omission of spatial lags of the systematic variables included in the model. The nuisance case refers to the occurrence of an autoregressive error structure as such, which may be the result of non-linearities, of spatial correlation among variables represented by the error term, or of a poor match between the spatial pattern of a phenomenon and the spatial cross-sections for which data is available. Below, both substantive spatial dependence and spatial dependence as a nuisance will be investigated, although non-linearities and a poor match are not considered as causes for the non-spherical structure of the error variance-covariance matrix.

3. Conventional remedies: the use of EGLS and ML estimators

The model given in (6)-(8) suggests that erroneously omitted, spatially lagged, systematic variables result in autoregressive residuals. Erroneously omitted, spatially lagged, systematic variables are a typical example of the omitted variable problem in a spatial context. This type of specification error causes the OLS estimator of the coefficients to be biased, with the bias being a linear combination of the coefficients of the omitted variables. Moreover, the true disturbance variance will, on average, be overestimated.

The conventional solution to the spatial dependence problem is not, however, a search for a proper respecification of the matrix X, but rather to assume that the true model is the model at hand and that the autocorrelation among the disturbances is due to spatial dependence as a nuisance. This approach and its consequences will be briefly described in this section.

The autoregressive structure is mostly represented by a stationary Markov scheme for the disturbances. Hence, in the general form for G orders of contiguity the model is specified as:

$$y = \mathbf{X}\beta + \varepsilon \qquad (9.1)$$

$$\varepsilon = \Sigma_g \lambda_g \mathbf{W}_g \varepsilon + \mu = \mathbf{W}_g^* \lambda + \mu \qquad (9.2)$$

where $\mathbf{W}_g$ is the weighting matrix corresponding to g-th order contiguity ($g = 1, 2, ..., G$), λ is a vector made up of the elements λ_g, with $|\lambda_g| < 1$ and $\Sigma_g \lambda_g < 1$, $\forall g$, in order to obtain a stationary process, $\mathbf{W}_g^*$ is the partitioned matrix $[\mathbf{W}_1 \varepsilon \mid \mathbf{W}_2 \varepsilon \mid ... \mid \mathbf{W}_G \varepsilon]$, and μ as before.

In accordance with the familiar time-series result the non-diagonal structure of the variance-covariance matrix, due to spatial dependence as a nuisance, causes the OLS estimators:

$$b_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' y \text{, and } s^2_{OLS} = [1/(R\text{-}k)]\, e'e \qquad (10)$$

to result in unbiased but inefficient parameter estimators, and a biased variance estimator.

The well-known EGLS time-series estimators due to Cochrane and Orcutt (1949) and Durbin (1960) have been extended to the spatial context [cf. Hordijk (1974), Ord (1975), Bartels (1979), Anselin (1980, 1981)]. The EGLS parameter estimator is given by:

$$b_{EGLS} = (\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}'\mathbf{A}y \qquad (11)$$

and the EGLS estimator for σ^2 is:

$$s^2_{EGLS} = [1/(R\text{-}k)]\, e'\mathbf{A}'\mathbf{A}\, e \qquad (12)$$

with $\mathbf{A} = \mathbf{I} - \Sigma_g \lambda_g \mathbf{W}_g$ and $e = \mathbf{y} - \mathbf{X}b_{EGLS}$. It should be observed that EGLS is only consistent when a consistent estimator is used for the nuisance parameters λ. If that is the case EGLS is not only formally equivalent to OLS applied to the suitably transformed variables $\mathbf{X}^* = \mathbf{A}\mathbf{X}$ and $y^* = \mathbf{A}y$, but also a maximum likelihood estimator.

The Cochrane-Orcutt (CO) procedure estimates λ by applying OLS to the auxiliary regression:

$$e = \Sigma_g \lambda_g \mathbf{W}_g e + \mu = \mathbf{W}_g^b \lambda + \mu \tag{13}$$

where e is the OLS estimated residual vector, $\mathbf{W}_g^b$ is the partitioned matrix $[\ \mathbf{W}_1 e \mid \mathbf{W}_2 e \mid ... \mid \mathbf{W}_G e\ ]$, and μ as before. The CO estimator can be applied in an iterative fashion by substituting the estimates $\hat{\lambda}$ in the matrix **A**, re-estimation of the vector of coefficients and σ^2 by means of (11) and (12), and so forth, until convergence. It is worth noticing that in spatial cross-section analysis data transformation does not lead to a loss of observations, so that the CO estimator is the analogue of the Prais-Winsten estimator frequently used in time-series analysis [cf. Judge et al. (1985, p. 286)]. When the nuisance parameters λ are estimated by OLS, however, the CO estimator is not consistent [cf. Anselin (1988)].

The Durbin estimator (DU) is obtained in a similar manner. Starting point is the common factor model:

$$y = \mathbf{X}\beta + \Sigma_g \lambda_g \mathbf{W}_g (y - \mathbf{X}\beta) + \mu$$

$$= \Sigma_g \lambda_g \mathbf{W}_g y + \mathbf{X}\beta - \Sigma_g \lambda_g \mathbf{W}_g \mathbf{X}\beta + \mu$$

$$= \mathbf{W}_g^c \lambda + \mathbf{X}\beta - \mathbf{W}_g^d \gamma + \mu \tag{14}$$

with $\mathbf{W}_g^c$ for $[\ \mathbf{W}_1 y \mid \mathbf{W}_2 y \mid ... \mid \mathbf{W}_G y\ ]$, $\mathbf{W}_g^d$ for $[\ \mathbf{W}_1 \mathbf{X} \mid \mathbf{W}_2 \mathbf{X} \mid ... \mid \mathbf{W}_G \mathbf{X}\ ]$, and $\gamma = (\lambda_1\beta', \lambda_2\beta', ..., \lambda_G\beta')'$. Model (14) can be estimated by OLS either in an unconstrained fashion or with the following $(G \times k)$ constraints imposed, $\hat{\beta}\hat{\lambda}_g = -\widehat{\beta\lambda_g}$, $\forall\, g$, with $\beta = (\beta_1, \beta_2, ..., \beta_k)'$. When a standardized weights matrix is used, the specification contains two constant terms which can of course not be identified separately. Consequently, $\mathbf{W}_g\mathbf{X}$ should be replaced by $\mathbf{W}_g\mathbf{X}^*$ and just $(G \times (k\text{-}1))$ constraints can be imposed. In practice, the DU estimator is mostly based on unconstrained estimation of (14) and subsequently (11) and (12) are estimated (non-iteratively) with $\hat{\lambda}$ taken from the regressors $\mathbf{W}_g^c$ in (14).

In contrast to the time-series case, the OLS estimator for λ in (13) and (14) is inconsistent irrespective of the properties of the error term μ [cf. Anselin (1988)]. Several ML procedures have been developed as alternatives to the estimators mentioned above [for an overview see Cliff and Ord (1973, 1981), Anselin (1980, 1981, 1988), Ripley (1981), Upton and Fingleton (1985), Griffith (1988)].

Assuming normally distributed errors, the ML estimator for the spatial autoregressive error model can be derived from the log-likelihood:

$$L = -(R/2)ln\pi - (R/2)ln\sigma^2 + ln|\mathbf{A}| - (1/2\sigma^2)(y\text{-}\mathbf{X}\beta)'\mathbf{A}'\mathbf{A}(y\text{-}\mathbf{X}\beta) \tag{15}$$

EGLS estimators for β and σ^2 are maximum likelihood estimators provided the estimator of the nuisance parameters λ are ML estimators. In that case ML estimates can be found through maximization of the concentrated log-likelihood conditional upon the estimates for e ($= y - \mathbf{X}\, b_{EGLS}$) as suggested in e.g. Brandsma and Ketellapper (1979a) and Anselin (1980). The problem is given by:

$$\max \quad L_C(\lambda) = -(R/2)ln\pi - (R/2)ln[(1/R)e'\mathbf{A}'\mathbf{A}e] + ln|\mathbf{A}| \tag{16.1}$$

$$\text{s.t.} \quad -1 < \lambda_g < 1 \,, \forall\, g \tag{16.2}$$

which results in estimates for λ. Via an iterative procedure with (11) and (12) MLERR estimates for the vector of unknowns $(\beta',\lambda',\sigma^2)'$ are obtained.

For higher order autoregressive error models with $\mathbf{W}_g$ not orthogonal the estimation procedure is numerically rather cumbersome [cf. Hepple (1976), Brandsma and Ketellapper (1979b)]. For $g = G = 1$ the spatial connectivity structure is obviously represented by a single weights matrix, and L_C is a univariate function. Solving for the autoregressive parameter λ_1 is then numerically not too complex, in particular as the Jacobian term $ln|\mathbf{A}|$ can be written as $\Sigma_i\, ln(1 - \lambda_1\, \omega_i)$ with ω_i for the eigenvalues of the spatial weights matrix $\mathbf{W}$ [cf. Ord (1975)]. See the appendix for details.

The following remarks apply. Although the maximum likelihood procedure has attractive asymptotic properties, it has been shown in Anselin (1980, 1981) that in small finite samples for the autoregressive error model the OLS and EGLS estimators may be superior in terms of bias, mean squared error, and mean absolute percentage error (see section five). It should be observed that in the simulations presented below the small finite sample properties of OLS and EGLS will not only be investigated in the case of the autoregressive error model but also in the context of substantive spatial dependence.

Secondly, in practice the estimators described above are pre-test estimators (*PTE*). In the case of EGLS and Moran's *IR* test the pre-test estimator is defined as [cf. e.g. King and Giles (1984)]:

$$b_{PTE} = P[IR > IR(\alpha)] \cdot b_{EGLS} + P[IR \leq IR(\alpha)] \cdot b_{OLS} \tag{17}$$

where $IR(\alpha)$ denotes the critical value of a 100α percent level test. Similar definitions apply to the ML estimator. The actual properties of the estimators presented here can markedly differ from their asymptotic properties because of the use of small samples and because of the pre-testing aspect (see also below).

4. The variable expansion method

The variable expansion method aims at remedying substantive spatial dependence by including erroneously omitted, spatially lagged variables into the set of explanatory variables.[1] Three kinds of spatial variable expansion methods will be elucidated below: limited spatial variable expansion (*LSVE*) which makes use of *LM* tests only; spatial variable expansion 1 (*SVE1*) which applies the autocorrelation test based on Moran's *IR* and *F* tests to the omitted, spatially lagged, explanatory variables; and spatial variable expansion 2 (*SVE2*) which applies *LM* autocorrelation tests, and *LR* or *F* tests to the omitted, spatially lagged, exogenous variables depending on whether or not the specification contains the spatially lagged, dependent variable.

In the limited spatial variable expansion strategy the only candidate for expansion is the spatially lagged dependent variable. The procedure is given in the following scheme.

---

1. Estimate the initial model $y = X\beta + \varepsilon$ by means of OLS.
2. Test the residuals by means of *LMERR* and *LMLAG*.
3. If neither leads to the rejection of H_0 of no spatial correlation among the residuals the ultimate model is the model obtained in step 1.
4. If *LMLAG* does not lead to the rejection of H_0 whereas *LMERR* does, or if both lead to the rejection of H_0 and the probability value (i.e. the probability that the sample value would be as large as the value actually observed if H_0 is true) corresponding to *LMERR* is smaller than the one corresponding to *LMLAG*, the spatial autoregressive error model is estimated by means of MLERR.
5. If *LMLAG* leads to the rejection of H_0, or if both lead to the rejection of H_0 and the probability value corresponding to *LMLAG* is smaller than the probability value corresponding to *LMERR* the mixed regressive spatial autoregressive model is estimated by MLLAG.

It should be observed that MLERR and MLLAG are ML estimators based on the assumption of spatially correlated errors and on the expansion of the initial model with a spatially lagged, dependent variable, respectively.

The *SVE1* strategy is a full expansion method based on *F* tests on spatially lagged, exogenous and spatially lagged, dependent variables. The *FOV* tests amount to applying the standard *F* test to the hypothesis $\gamma = 0$ in the augmented regression:

$$y = X\beta + \gamma z + \varepsilon \tag{18}$$

where z is a vector representing an omitted spatially lagged variable. The null hypothesis is tested with an F test, which is formulated in terms of the residual sum of squares of the restricted and the unrestricted model:

$$FOV = \frac{(e_R' e_R - e_U' e_U) / q}{e_U' e_U / (R - k)} \sim F(q, R\text{-}k) \tag{19}$$

where e_R' and e_U' are the estimated vectors of errors of the restricted and unrestricted models, respectively, and q is the number of variables omitted in the restricted model. It should be observed that when the model has been augmented by the spatially lagged dependent variable the F test is inappropriate. Instead the Likelihood Ratio (*LR*) test should be applied. For the following reasons, however, the F test will be used here. First, the asymptotic properties of the *LR* test may not hold in small samples (which will be analyzed in the simulations). Secondly, the F test is likely to be better known than the *LR* test, and may therefore be more easily applicable in practice.

The *SVE1* strategy can be summarized by means of the steps given in the scheme below.

1. Estimate the initial model $y = \mathbf{X}\beta + \varepsilon$ by means of OLS.
2. Identify the subset S of systematic explanatory variables for which the inclusion as spatially lagged variables into the model is plausible on theoretical grounds.
3. Expand the initial model by successively including variables from S (one at a time) and estimate the models $y = \mathbf{X}\beta + \rho_{i,g} \mathbf{W}_g z_i + \varepsilon$, $z_i \in S$, $\forall\, i,g$.
4. Test the hypothesis $\rho_{i,g} = 0$, $z_i \in S$, $\forall\, i,g$, by means of *FOV*.
5. Let E be the set of coefficients for which the hypothesis $\rho_{i,g} = 0$ is rejected. Include z_i corresponding to $\rho_{i,g} \in E$ with the smallest probability value into the model.
6. Steps 4. and 5. are repeated for the remaining relevant variables and orders of contiguity.
7. The residuals of the ultimate model are calculated and tested for spatial dependence by means of *IR*. If the hypothesis of no spatial dependence is rejected, the MLERR estimator is applied to the initial model.

The following remarks apply. First, in step 7, OLS residuals are used to test for spatial correlation among the true errors. The OLS residuals are correlated whether or not the unobserved population errors are correlated. Simulation studies by Bartels and Hordijk (1977) and Brandsma and Ketellapper (1979a) show, however, that the power of *IR* is stronger for OLS-residuals than for a residual estimator with a scalar covariance matrix (such as BLUS, LUS, and RELUS). Secondly, application of MLERR, in step 7, to a model containing the spatially lagged dependent variable is not feasible because of the identification problems related to the use of just one weights matrix for both substantive and nuisance dependence. Hence, one has to resort to MLERR estimation of the initial model. In the case of the presence of both spatially lagged exogenous variables and autoregressive errors MLERR would be feasible. As the latter model is not one of the generating models it will seldom occur, and has therefore been disregarded here.

The *SVE1* procedure is based on generally well-known tests and the OLS estimator. It is therefore easy to apply, and likely to be popular in applied research. However, when the unrestricted model contains the spatially lagged endogenous variable the OLS estimator is inconsistent and biased. In that case the MLLAG estimator should be used instead. Moreover, in the model containing the spatially lagged endogenous variable the standard *F* test should be replaced by the *LR* test [cf. Anselin (1988)].

Strategy *SVE2* is a full expansion method based on *LMERR* and *LMLAG* tests on the residuals, and *LR* or *FOV* tests on the coefficients of the spatially lagged explanatory variable. In the present procedure the methodological flaws of strategy *SVE1* are avoided. It should be observed, however, that the features of the present procedure are based on asymptotic theory, whereas in practice usually small samples are used. The procedure is made up of the following steps.

---

1. Estimate the initial model $y = X\beta + \varepsilon$ by means of OLS.
2. Identify the subset S of systematic explanatory variables for which the inclusion as spatially lagged variables into the model is plausible on theoretical grounds.
3. Test the hypotheses of no residual nor substantive spatial autocorrelation by means of *LMERR* and *LMLAG*.
4. If both tests lead to acceptance of the null hypotheses test the hypothesis $\rho_{i,g} = 0$, $z_i \in S^*$ where S^* is the subset of S containing the exogenous variables only, $\forall\ i,g$, by means of *FOV*. Include z_i corresponding to $\rho_{i,g} \in E$ (i.e. the set of coefficients found to be unequal to zero) with the smallest probability value into the model. Repeat this step for the remaining relevant variables and orders of contiguity.
5. If *LMERR* is significant and *LMLAG* not, or if both tests lead to the rejection of H_0 and the probability value corresponding to the former is smaller than the one corresponding to the latter, the model with autoregressive errors is taken as the ultimate model and is estimated by means of MLERR.
6. If *LMLAG* is significant and *LMERR* not, or if both tests lead to the rejection of H_0 and the probability value corresponding to the former is smaller than the one corresponding to the latter, the mixed regressive spatial autoregressive model is estimated by means of the MLLAG estimator. Next, the model is successively expanded by spatially lagged exogenous variables following the lines given in step 4, although with *LR* tests instead of *FOV* tests.

The following remarks apply to the variable expansion methods described above. First, these procedures are the opposite of the classical way of model selection, which consists of the estimation of a tentative, general model and the deletion of those variables for which the estimated coefficients either have a wrong sign or are not significantly different from zero.

Secondly, a possible alternative to the variable expansion method is to start with a general model including the various plausible, spatially lagged variables and to reduce the model in the classical way [cf. Hendry and Richard (1982)]. The advantages of the variable expansion methods described earlier are the, at least initially, lower sensitivity to multicollinearity and greater parsimony with respect to degrees of freedom. On the other hand, the variable expansion method is basically a forward step-wise regression. Exclusion of relevant variables does not only lead to biased, inconsistent estimators of the regression coefficients but the true disturbance variance will also be overestimated. The Hendry approach, which is essentially a backward step-wise regression, does not suffer from inconsistency or bias. In this paper the properties of forward step-wise methods have been investigated. An interesting and important topic for further research would be a comparison of these procedures with the Hendry approach.

Thirdly, the models selected by means of the variable expansion methods clearly have a data-instigated nature. That is, the data set which has been used to find a model that fits the data at hand is also used to estimate the variability of the estimator. Because the data gave birth to it, the accuracy of the estimator of the parameter vector of a data-instigated model will be over-estimated to an unknown extent, and the goodness of fit to the sample data is likely to be greater than the fit to the population [cf. e.g. Leamer (1978), Lovell (1983), Folmer (1988) with respect to autocorrelation in time series models]. The variable expansion method should primarily be applied in exploratory analyses. Moreover, only those spatially lagged explanatory variables should be considered for inclusion into the model for which there exists sufficient theoretical justification. Insight into the robustness of the model obtained via variable expansion methods (and other kinds of data-instigated models) can be attained by means of cross validation [cf. e.g. Mosteller and Tukey (1977) for details]. The problem of data-instigated models will not be considered any further here. For information about these problems, which frequently occur in the context of any kind of applied econometric research, the reader is referred to, among others, Judge and Bock (1978), Hendry and Richard (1982), *Auctores varii* (1984), and Gilbert (1986).

5. Experimental design and simulation results

In this section the small sample performance of the variable expansion methods and of the various estimators mentioned in section three will be investigated in the case of both substantive spatial correlation and spatial dependence as a nuisance. The model chosen for experimentation reads as:

$$y = \zeta \mathbf{W} y + \mathbf{X} \beta + \mathbf{W} \mathbf{X}^{*} \rho + (\mathbf{I} - \lambda \mathbf{W})^{-1} \mu \tag{21}$$

where y is the $(R \times 1)$ dependent variate vector with $R = 26$; $\mathbf{X}$ is made up of x_0, the unit vector, and of x_1 and x_2 which are drawings from a uniform (0,10) distribution; $\mathbf{W}$ is the standardized binary first-order contiguity matrix of the O'Sullivan data for Ireland given in unstandardized form in Cliff and Ord (1981, p. 230); and $\mu = (\mu_1, \mu_2, ..., \mu_{26})' \sim N(0, \sigma_\mu^2 \mathbf{I})$ where $\mathbf{I}$ is the $(R \times R)$ identity matrix and $\sigma_\mu^2 = 2.0$. The true values contained in the vector β are all set to 0.5. The elements of $\mathbf{X}$ are fixed between the experiments, and the characteristics of x_1 and x_2 are appropriate in terms of e.g. multicollinearity (means 4.003 and 5.157, standard deviations 2.815 and 2.529, and $r_{x_1 x_2} = 0.156$). The error variance has been set to 2.0 causing the estimated R^2 to be on average approximately 0.55 when $\rho = (0, 0, ..., 0)'$ and $\zeta = \lambda = 0$, which seems to be a realistic bench-mark value found in applied spatial research. As mentioned in section two, different weights matrices are required for reasons of identification if the complete model were to be estimated. In the simulations, however, only genuine submodels are analyzed, which makes it possible to use a single weights matrix.

The following model is estimated by OLS:

$$y = \mathbf{X} \beta + \varepsilon \tag{22}$$

where ε is assumed to be white noise, whereas the true models are the regressive-spatial crossregressive, the regressive-spatial autoregressive, the regressive-spatial regressive and the autoregressive error model, respectively. Since in some runs the coefficients of the spatially lagged variables and the nuisance or autocorrelation parameter are zero, model (22) is also one of the true models.

For the EGLS and ML estimators the null-hypothesis of no (substantive) spatial autocorrelation has been tested by means of the Moran coefficient with a nominal (two-sided) Type-I error of 0.05. The variable expansion methods are sequential. This means that the nominal Type-I errors of the individual test should be adjusted so as to obtain the desired nominal α-level of 0.05 for the procedure as a whole. As the number of tests is not known *a priori*, a correction for multiple comparisons, such as Bonferroni bounds, cannot be used [cf. Savin (1980)]. However, assuming independence of the individual tests upper and lower bounds of the overall level can be derived on the basis of the minimum and maximum number of comparisons required. Levels for the individual tests and overall levels for the various conventional and expansion procedures are given in Table 1.

Table 1: α-levels for individual tests and overall α-levels for the different test strategies.*

Method/ tests	Conventional method	Limited spatial variable expansion	Full spatial variable expansion variant 1	Full spatial variable expansion variant 2
IR	0.05		0.01	
LMERR		0.025		0.01
LMLAG		0.025		0.01
FOV			0.01	0.01
LR				0.01
Minimum # comparisons	1	2	4	2
Maximum # comparisons	1	2	7	5
Overall α-level	0.05	0.05	0.04-0.07	0.02-0.05

* For the approximate overall α-levels it is assumed that the tests are independent.

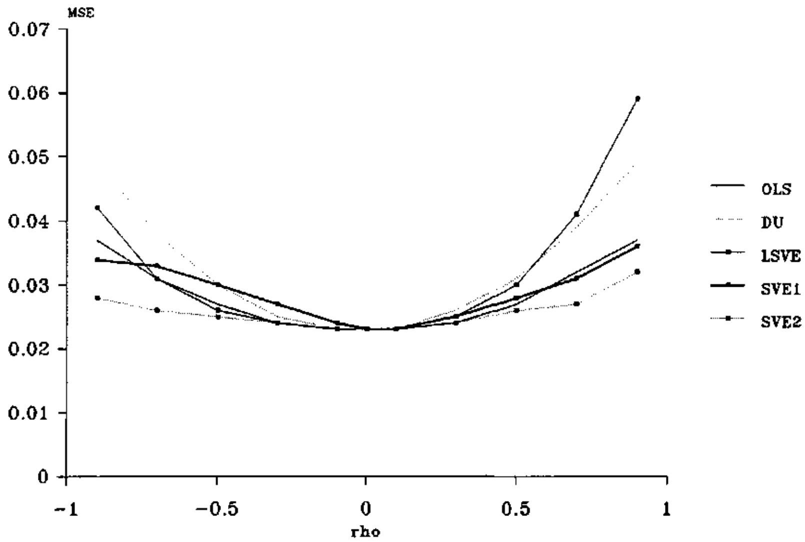
The overall results are presented in terms of power of the tests, probabilities of finding the true model, absolute parameter bias (BIAS), mean squared error (MSE), and mean absolute percentage error (MAPE). The fit indices have been averaged over the number of samples. In addition, the fit indices regarding the β coefficients of the x variables (excluding the constant term) have been averaged over the number of regressors also. The same applies to the coefficients of the spatially lagged x variables. The following definitions apply:

$$\text{BIAS} = [1/(m \times N)] \Sigma_i \Sigma_j |b_{ij} - \beta_{ij}| \tag{23.1}$$

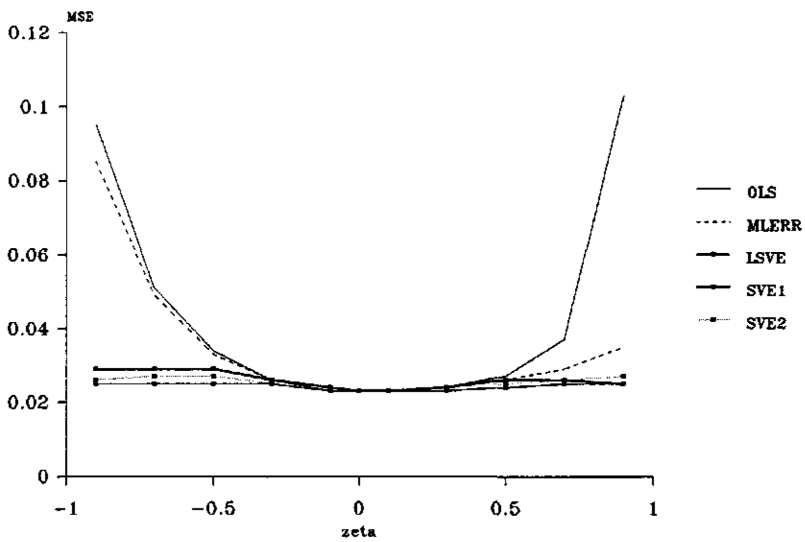
$$\text{MSE} = [1/(m \times N)] \Sigma_i \Sigma_j (b_{ij} - \beta_{ij})^2 \tag{23.2}$$

$$\text{MAPE} = [1/(m \times N)] \Sigma_i \Sigma_j |(\beta_{ij} - b_{ij}) / \beta_{ij}| \tag{23.3}$$

where i (= 1, 2, ..., m) indexes the (spatially lagged) x variables, and j (= 1, 2, ..., 5,000) the samples.

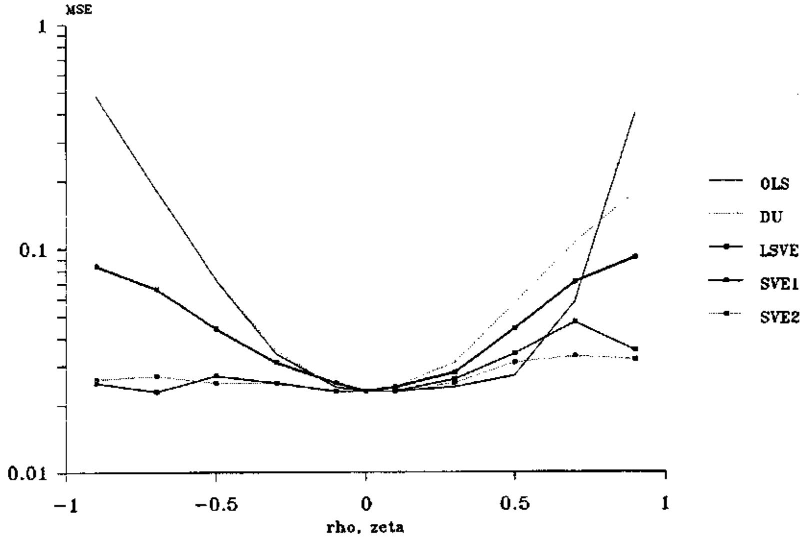


Crossregressive model

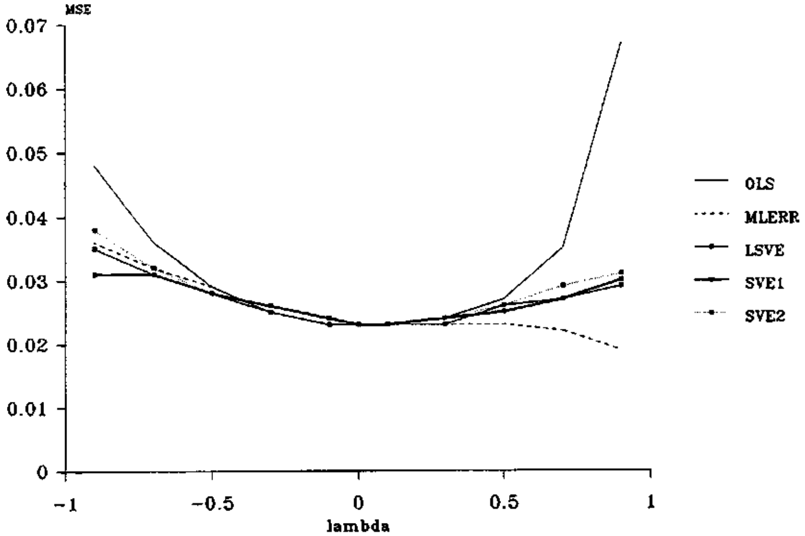


Autoregressive model

Figure 1: Mean squared error for OLS, CO or DU or MLERR, and the spatial variable expansion methods *LSVE*, *SVE1*, and *SVE2* for four different true models. The fit index is averaged over both the number of samples and the β coefficients. An irregularly spaced vertical axis implies that a log scale is used, and all results except for OLS refer to pre-test estimators.

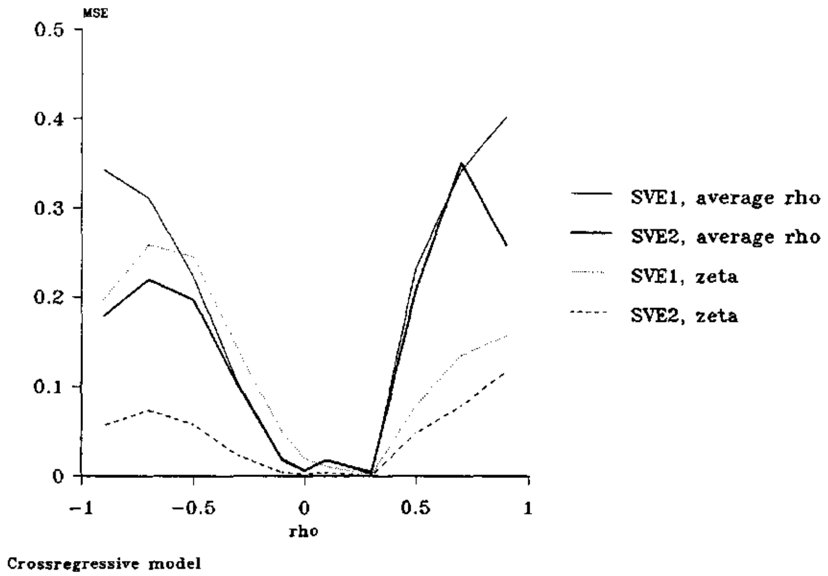


Spatial regressive model



Autoregressive error model

Figure 1: continued.



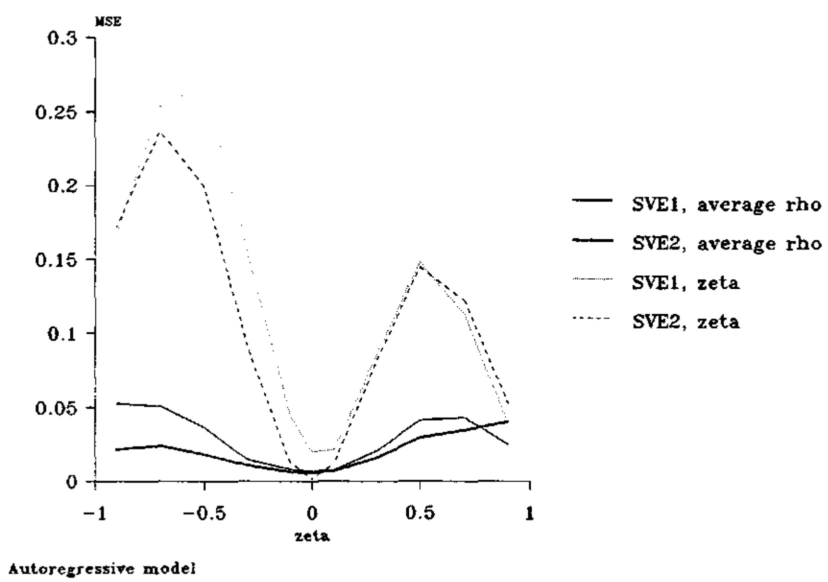
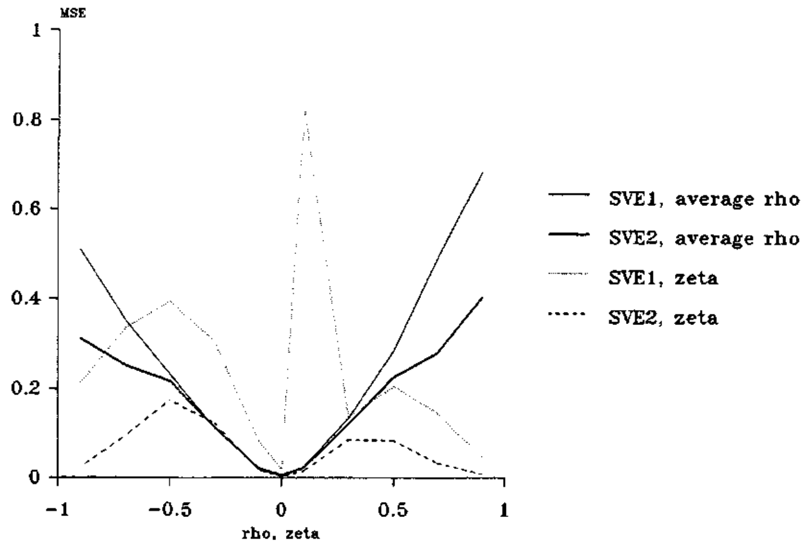
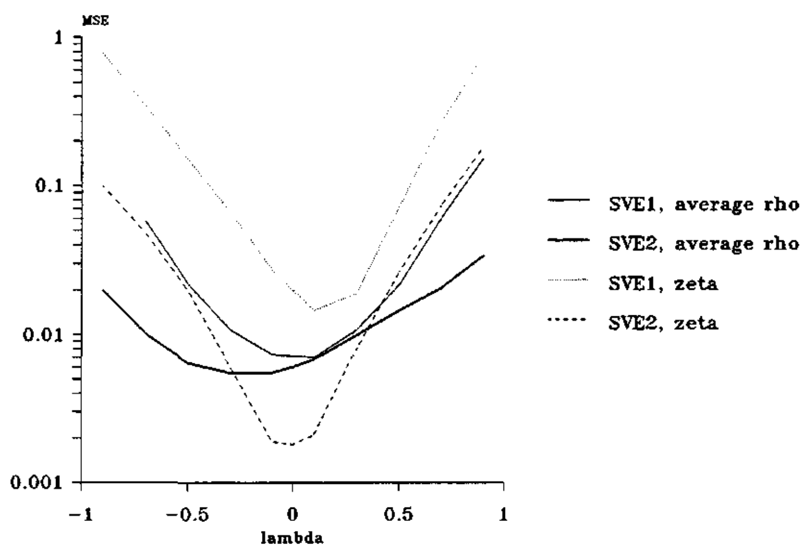


Figure 2: Mean squared error for the coefficients of the spatially lagged, exogenous variables (ρ) and the spatially lagged, dependent variable (ζ) for the spatial variable expansion methods *SVE1* and *SVE2*. The fit index is averaged over the number of samples, and also over the ρ's. An irregularly spaced vertical axis implies that a log scale is used.



Spatial regressive model



Autoregressive error model

Figure 2: continued

Table 2: Probability of rejecting the null hypothesis of no (substantive) spatial dependence for four different true models.[a]

Parameter	True model											
	Crossregressive			Autoregressive			Spatial regressive			Autoregressive errors		
	IR	*LMERR*	*LMLAG*	*IR*	*LMERR*	*LMLAG*	*IR*	*LMERR*	*LMLAG*	*IR*	*LMERR*	*LMLAG*
-0.9	0.297	0.080	0.223	0.222	0.144	0.890	0.003	0.001	0.996	0.482	0.358	0.312
-0.7	0.206	0.052	0.260	0.162	0.096	0.701	0.011	0.001	0.960	0.288	0.193	0.184
-0.5	0.124	0.030	0.221	0.105	0.058	0.401	0.026	0.058	0.819	0.147	0.084	0.101
-0.3	0.067	0.017	0.121	0.056	0.056	0.148	0.042	0.012	0.469	0.063	0.033	0.051
-0.1	0.043	0.011	0.035	0.033	0.013	0.036	0.036	0.013	0.072	0.032	0.012	0.022
0	0.040	0.011	0.021	0.040	0.011	0.021	0.040	0.011	0.021	0.040	0.011	0.021
0.1	0.040	0.010	0.032	0.060	0.015	0.038	0.069	0.018	0.085	0.060	0.014	0.028
0.3	0.066	0.015	0.140	0.195	0.063	0.201	0.362	0.151	0.600	0.170	0.053	0.074
0.5	0.120	0.029	0.319	0.483	0.253	0.540	0.810	0.576	0.952	0.405	0.191	0.188
0.7	0.205	0.052	0.474	0.802	0.615	0.836	0.990	0.948	0.999	0.692	0.472	0.410
0.9	0.288	0.082	0.574	0.969	0.913	0.970	1.000	1.000	1.000	0.918	0.814	0.733

[a] Although not applicable here the adjustments necessary in the context of comparing the different estimators make that the nominal Type-I errors differ between the tests. For Moran's *IR* test the nominal Type-I error is 0.05, whereas for the *LM* tests it is set to 0.025.

Table 3: Probability of finding the true model for four different true models.

Parameter	True model											
	Crossregressive			Autoregressive			Spatial regressive			Autoregressive errors		
	LSVE	*SVE1*	*SVE2*	*LSVE*	*SVE1*	*SVE2*	*LSVE*	*SVE1*	*SVE2*	*LSVE*	*SVE1*	*SVE2*
-0.9	--	0.315	0.825	0.884	0.849	0.772	--	0.071	0.466	0.241	0.190	0.115
-0.7	--	0.110	0.709	0.692	0.714	0.526	--	0.012	0.492	0.144	0.151	0.050
-0.5	--	0.020	0.451	0.397	0.434	0.251	--	0.000	0.186	0.064	0.100	0.018
-0.3	--	0.003	0.158	0.144	0.159	0.079	--	0.000	0.024	0.026	0.049	0.005
-0.1	--	0.000	0.032	0.034	0.038	0.011	--	0.000	0.001	0.009	0.028	0.002
0	0.972	0.926	0.971	0.972	0.926	0.971	0.972	0.926	0.971	0.972	0.926	0.971
0.1	--	0.000	0.026	0.037	0.017	0.019	--	0.000	0.004	0.010	0.052	0.004
0.3	--	0.002	0.140	0.192	0.102	0.126	--	0.001	0.079	0.035	0.136	0.018
0.5	--	0.021	0.370	0.516	0.400	0.406	--	0.000	0.329	0.124	0.272	0.079
0.7	--	0.111	0.544	0.799	0.752	0.701	--	0.012	0.506	0.291	0.305	0.228
0.9	--	0.306	0.564	0.934	0.932	0.853	--	0.072	0.383	0.453	0.151	0.417

Table 4: The best estimator in terms of bias, mean squared error, and mean absolute percentage error for the coefficients of the *x* variables (excluding the constant term) for four different true models.[a]

Parameter	True model											
	Crossregressive			Autoregressive			Spatial regressive			Autoregressive errors		
	BIAS	MSE	MAPE	BIAS	MSE	MAPE	BIAS	MSE	MAPE	BIAS	MSE	MAPE
-0.9	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	SVE1	SVE1	SVE1
-0.7	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	SVE1	SVE1	SVE1
-0.5	SVE2	SVE2	SVE1	LSVE	LSVE	LSVE	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE
-0.3	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE
-0.1	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	LSVE	SVE2	OLS	SVE2
0	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
0.1	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
0.3	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE	OLS	OLS	OLS	ML	ML	ML
0.5	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE	OLS	OLS	OLS	ML	ML	ML
0.7	SVE2	SVE2	SVE2	LSVE	LSVE	LSVE	SVE2	SVE2	SVE2	ML	ML	ML
0.9	SVE2	SVE2	SVE2	SVE1	LSVE	SVE1	SVE2	SVE2	SVE2	ML	ML	ML

[a] All estimators except OLS are pre-test estimators.

The following additional technical details are relevant. The number of replications (or samples) has been set to 5000. The DU estimator has been applied in an unconstrained manner. In order to limit the number of variations in the simulation experiments, no distinction in parameter value has been made with regard to the elements of the vector $\rho = (\rho_1\ \rho_2)'$, i.e. $\rho_1 = \rho_2$. Regarding the regressive-spatial regressive model the same restriction applies, i.e. $\zeta = \rho_1 = \rho_2$. Finally, for the same reason the simulation experiments have been limited to first-order contiguity which restricts the number of possible true models considerably. The latter does not necessarily imply that the results easily generalize to the higher order contiguity cases. From a technical point of view the use of higher order contiguities makes that the likelihood function becomes multivariate in terms of the spatial parameters to be estimated (see section three). Moreover, variants of the *LMLAG* statistic have to be developed in order to test for omitted, g-th order spatially lagged, dependent variables when the specification already contains the g'-th order, spatially lagged, dependent variable ($g \neq g'$).

The overall results for the conventional and variable expansion methods are summarized in the Tables 2 - 4 and in the Figures 1 and 2. Table 2 gives an overview of the power of Moran's *IR* test and of the *LM* tests for four different true models. Table 3 gives the probability of finding the true model for the spatial variable expansion methods. Figure 1 refers to the MSE of the β coefficients of OLS, the variable expansion methods, and the best estimator of the conventional estimators CO, DU and ML, whereas Figure 2 gives the MSE for the coefficients of spatially lagged variables of the full spatial variable expansion methods. Finally, Table 4 gives an overview of the best estimator of the β coefficients in terms of BIAS, MSE, and MAPE.

From Table 2 it follows that when the true model is the mixed regressive-spatial crossregressive model the power of *LMERR* and of the test based on *IR* is highly insufficient. *LMLAG* outperforms its alternatives, although it is still insufficient. In the case of a mixed regressive-spatial autoregressive model *LMLAG* is superior, although *LMERR* and the *IR* based test perform reasonably well for positive values of ζ. In the case of autoregressive errors the *IR* based test outperforms *LMLAG* and also *LMERR*. However, for negative values of λ the power of all three tests is insufficient. Finally, *LMLAG* is superior in the mixed regressive-spatial regressive model. The results found here are in conformity with various other studies, in particular Anselin and Rey (1991), and Florax and Folmer (1991).

Table 3 gives the probabilities of finding the true model. In the case of a mixed regressive-spatial crossregressive generating model the variable expansion method based on Lagrange Multipliers and Likelihood ratio tests (*SVE2*) outperforms the method based on Moran's *IR* and *FOV* tests (*SVE1*), although for positive values of ρ the performance is insufficient. For the mixed regressive-spatial autoregressive model *SVE1* outperforms *SVE2*. Moreover, for positive values of ζ it is slightly inferior to the combination of LMERR and LMLAG tests (*LSVE*), which is tailored to this problem. In the cases of autoregressive errors and the mixed regressive spatial regressive model all three procedures perform rather poorly.

In Figure 1 the mean squared error is depicted for the four different true models. From this figure the conclusion can be drawn that *SVE2* usually belongs to the procedures with lowest MSE. A major exception is the autoregressive error model where Moran's *IR* test followed by CO, DU (not in the figure) or ML estimation of a model with autoregressive errors is superior. The difference between *SVE2* on the one hand and CO, DU and ML on the other, however, is not substantial. Another major conclusion that follows from Figure 1 is that all three expansion methods perform about equally well in terms of MSE. In particular, the size of their MSE seldom exceeds 10 percent of the true value.

Figure 2 shows the MSE of the estimators of the coefficients of the spatially lagged variables for four different true models. It can be concluded that *SVE2* outperforms *SVE1*, though mostly only slightly except for the spatial regressive model with low positive values of ρ and ζ. A major exception of the superiority of *SVE2* is the autoregressive model with values of $\zeta > 0.5$. It should be observed that the size of the MSE for the coefficients of the spatially lagged variables is substantially larger than in the case of the β coefficients. Given the low probability of finding the true model in many instances (see Table 3) this is not surprising. In many situations in which the spatial variable expansion methods are applied, the interest is not primarily on the coefficients of the spatially lagged variables or the nuisance parameters. The latter are mainly taken into consideration in order to obtain adequate estimators of the coefficients of the variables initially included into the model. Therefore, one may conclude from Figures 1 and 2 that the spatial variable expansion methods perform satisfactorily.

Similar results hold in terms of bias, variance, and mean absolute percentage error. They are not presented in detail here for reasons of space [cf. Florax (1992) for more detailed results]. An overview of the performance in terms of BIAS, MSE, and MAPE is given in Table 4. From this table it follows that there is no uniformly superior procedure. However, *SVE2* outperforms the other procedures in the cases of the crossregressive and the spatial regressive model. For the autoregressive and the autoregressive error model *LSVE* and the EGLS and ML estimators are superior respectively, although the differences with respect to the full variable expansion methods are relatively small. *SVE1* is in most cases only slightly inferior to *SVE2*.

6. Conclusions

At the outset of this paper it has been observed that spatial autocorrelation among the residuals of linear spatial models is a serious problem in applied spatial research. The conventional solution to this problem is to assume that the true model is the initial model with an autoregressive error structure. Estimators, such as the inconsistent CO or DU estimator and the consistent ML estimator are subsequently applied.

From this paper the overall conclusion emerges that spatial variable expansion methods outperform the conventional procedures. This applies in particular to the spatial variable expansion method based on *LM* and *LR* tests (*SVE2*). However, in various cases the expansion methods do not identify the correct model. This may be due to the individual α-levels, which were set relatively low so as to make the overall α-level of the conventional and the expansion methods about the same size. In practical applications it is preferable to use higher values for α, because the consequences of erroneously assuming substantive correlation has less serious consequences than erroneously accepting the hypothesis of no substantive spatial correlation.

It should be noted that the results are likely to be dependent upon the specification of the weights matrix and the size of the spatial system in terms of the number of regions. So, further investigations are needed for both regular and irregular lattice structures, and for various sample sizes. Moreover, differences over space may also show up in spatial heterogeneity. The joint occurrence of heterogeneity and dependence over space has not been investigated here, but may also have implications for the performance of the different approaches.

Appendix

The maximization of the concentrated log-likelihood given in (16.1-2) corresponds to the minimization of the function $F(\lambda)$, formally:

$$\min \quad F(\lambda) = (R/2)\, ln[(1/R)e'\mathbf{A}'\mathbf{A}e] - ln|\mathbf{A}|$$
$$\text{s.t.} \quad -1 < \lambda < 1$$

with $\mathbf{A} = \mathbf{I} - \lambda\mathbf{W}$. A solution to the first-order condition $dF(\lambda)/d\lambda = 0$ can be obtained in a straightforward manner and reads as [cf. also Hepple (1976), Brandsma and Ketellapper (1979a)]:

$$d\,F(\lambda)/d\lambda = -R\,(e'\mathbf{A}'\mathbf{W}e \,/\, e'\mathbf{A}'\mathbf{A}e) + \mathrm{tr}(\mathbf{A}^{-1}\mathbf{W}) = 0 \qquad (A.1)$$

The minimization problem can be easily solved using a bisection routine, e.g. as given in Anselin (1988, pp. 216-217). Prefabricated routines, based on for instance steepest descent [cf. the UVMID routine in IMSL (1987, pp. 795-798)], work on the basis of adjustments per iteration defined in terms of the derivative of the log likelihood, a procedure that may easily result in out of bound solutions.

In order to avoid multiple computations of the determinant the Jacobian term can be rewritten prior to the minimization of $F(\lambda)$ as a sum of eigenvalues, that is:

$$ln|\mathbf{A}| = \Sigma_{i=1}^{R}\, ln(1 - \lambda\omega_i) \qquad (A.2)$$

with ω_i for the eigenvalues of $\mathbf{W}$. In a FORTRAN programming environment the eigenvalues can be computed with IMSL routine EVLRG [cf. IMSL (1987, pp. 293-294)].

For symmetric $\mathbf{W}$ all eigenvalues are real and Ord (1975) contains some useful analytical results [cf. also Griffith (1988)]. Ord (1975) also presents a procedure to transform asymmetric $\mathbf{W}$, which may have one or more complex eigenvalues, into a symmetric block diagonal form. The following procedure is an attractive alternative based on the original weighting matrix, regardless whether it is symmetric or not. The procedure is computational efficient and avoids complex number arithmetics. For any real $(R \times R)$ weighting matrix $\mathbf{W}$ with real and/or complex eigenvalues, and for all λ with $|\mathbf{I} - \lambda\,\mathbf{W}| \geq 0$:

$$ln|\mathbf{I} - \lambda\mathbf{W}| =$$
$$ln[(1-\lambda\omega_1)(1-\lambda\omega_2)\ \ldots\ (1-\lambda\omega_R)] =$$
$$ln[\,|1-\lambda\omega_1|\ \ldots\ |1-\lambda\omega_R|\,] =$$
$$\Sigma_{i=1}^{R}\, ln\,|1-\lambda\omega_i| =$$
$$\Sigma_{i=1}^{R}\, ln\,\sqrt{(\mathrm{Re}(1-\lambda\omega_i))^2 + (\mathrm{Im}(1-\lambda\omega_i))^2} =$$
$$\tfrac{1}{2}\,\Sigma_{i=1}^{R}\, ln[(1-\lambda\mathrm{Re}\omega_i)^2 + \lambda^2(\mathrm{Im}\omega_i)^2] \qquad (A.3)$$

Straightforward application of calculus gives the first derivative of (A.3) as:

$$\Sigma_{i=1}^{R}\, \{[(\lambda\mathrm{Re}\omega_i-1)\mathrm{Re}\omega_i + (\mathrm{Im}\omega_i\lambda)^2] \,/\, [(1-\lambda\mathrm{Re}\omega_i)^2 + \lambda^2(\mathrm{Im}\omega_i)^2]\} \qquad (A.4)$$

Notes

[1] It should be observed that the variable expansion method bears some similarity to the spatial (parameter) expansion method [Casetti (1972, 1991)]. This approach boils down to the following procedure. Assume the following simple linear regression model with one explanatory variable (i.e. the initial or restricted model):

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad (1')$$

Spatial dependence may *inter alia* show up when the regression coefficients are subject to parameter drift over the spatial observation units. The spatial expansion approach amounts to the assumption that the spatial variation in the parameters can be described as an exact function of a finite number, say E, of expansion variables:

$$\beta_1 = \gamma_0 + \gamma_1 s_{1,r} + \gamma_2 s_{2,r} + \ldots + \gamma_E s_{E,r} \, , \forall \, r \qquad (2')$$

where $s_{e,r}$ (e = 1, 2, ..., E) are indexes denoting a region's position in the spatial system, mostly specified in terms of trend surface polynomials on the basis of location coordinates, or orthogonal principal components of trend surface expansions. After substitution of the expanded parameter and rewriting, equation (2´) with all regions of the spatial system included, reads as:

$$y = \beta_0 + \gamma_0 x + \gamma_1 \mathbf{S}_1 x + \gamma_2 \mathbf{S}_2 x + \ldots + \gamma_E \mathbf{S}_E x + \varepsilon \qquad (3')$$

where $\mathbf{S}_e = \text{diag}(s_{e,r})$, i.e.:

$$\mathbf{S}_e = \begin{cases} s_{i,j} = s_{e,r} & \text{if } i = j \\ s_{i,j} = 0 & \text{otherwise.} \end{cases} \qquad (4')$$

As has been shown theoretically by Anselin (1988, pp. 127-129) and empirically by Jones (1983) and Casetti and Jones (1988) spatial parameter expansion may result in reduced spatial dependence as measured by the residual Moran coefficient. However, instead of spatial expansion of the parameter(s), which is a practical and attractive remedy for spatial heterogeneity, a different interpretation should be given to spatial expansion when spatial dependence occurs.

References

Anselin, L., *Estimation Methods for Spatial Autoregressive Structures*, Ithaca, New York: Cornell University, Regional Science Dissertation and Monograph Series #8, 1980.

Anselin, L., Small Sample Properties of Estimators for the Linear Model with a Spatial Autoregressive Structure in the Disturbance, *Modeling and Simulation*, vol. 12, pp. 899-904, 1981.

Anselin, L., A Note on Small Sample Properties of Estimators in a First-Order Spatial Autoregressive Model, *Environment and Planning A*, vol. 14, pp. 1023-1030, 1982.

Anselin, L., *Specification Tests and Model Selection for Spatial Interaction and the Structure of Spatial Dependence*, Ohio: Ohio State University, Department of City and Regional Planning, 1985.

Anselin, L., *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers, 1988.

Anselin, L., Spatial Dependence and Spatial Structural Instability in Applied Regression Analysis, *Journal of Regional Science*, vol. 30, pp. 185-207, 1990.

Anselin, L. and A.D. Griffith, Do Spatial Effects Really Matter in Regression Analysis?, *Papers of the Regional Science Association*, vol. 65, pp. 11-34, 1988.

Anselin, L. and S. Rey, Properties of Tests for Spatial Dependence in Linear Regression Models, *Geographical Analysis*, vol. 23, pp. 112-131, 1991.

Auctores varii, Journal of Econometrics, vol. 25, 1984.

Bartels, C., Operational Statistical Methods for Analysing Spatial Data, in: C. Bartels, and R. Ketellapper (eds), *Exploratory and Explanatory Analysis of Spatial Data*, Boston: Martinus Nijhoff, 1979, pp. 5-50.

Bartels, C.P.A. and L. Hordijk, On the Power of the Generalized Moran Contiguity Coefficient in Testing for Spatial Autocorrelation Among Regression Disturbances, *Regional Science and Urban Economics*, vol. 7, pp. 83-101, 1977.

Brandsma, A. and R. Ketellapper, Further Evidence on Alternative Procedures for Testing of Spatial Correlation Among Regression Disturbances, in: C. Bartels, and R. Ketellapper (eds), *Exploratory and Explanatory Analysis of Spatial Data*, Boston: Martinus Nijhoff, 1979a, pp. 111-136.

Brandsma, A. and R. Ketellapper, A Biparametric Approach to Spatial Autocorrelation, *Environment and Planning A*, vol. 11, pp. 51-58, 1979b.

Burridge, P., Testing for a Common Factor in a Spatial Autoregressive Model, *Environment and Planning A*, vol. 13, pp. 795-800, 1981.

Casetti, E., Generating Models by the Expansion Method: Applications to Geographical Research, *Geographical Analysis*, vol. 4, pp. 81-91, 1972.

Casetti, E., The Investigation of Parameter Drift by Expanded Regressions: Generalities, and a 'Family Planning' Example, *Environment and Planning A*, vol. 23, 1991.

Casetti, E. and J.-P. Jones, Spatial Parameter Variation by Orthogonal Trend Surface Expansions: An Application to the Analysis of Welfare Program Participation Rates, *Social Science Research*, vol. 16, pp. 76-88, 1988.

Cliff, A.D. and J.K. Ord, Testing for Spatial Autocorrelation Among Regression Residuals, *Geographical Analysis*, vol. 4, pp. 267-284, 1972.

Cliff, A.D. and J.K. Ord, *Spatial Autocorrelation*, London: Pion, 1973.

Cliff, A.D. and J.K. Ord, *Spatial Processes, Models & Applications*, London: Pion, 1981.

Cochrane, D. and G. Orcutt, Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms, *Journal of the American Statistical Association*, vol. 44, pp. 32-61, 1949.

Doreian, P., Linear Models with Spatially Distributed Data: Spatial Disturbances or Spatial Effects?, *Sociological Methods and Research*, vol. 9, pp. 29-60, 1980.

Durbin, J., Estimation of Parameters in Time-Series Regression Models, *Journal of the Royal Statistical Society B*, vol. 22, pp. 139-153, 1960.

Florax, R., *The University: A Regional Booster? Economic Impacts of Public Knowledge Infrastructure*, Aldershot: Avebury, 1992 (forthcoming).
Florax, R. and H. Folmer, *Robust Testing for Spatial Correlation among Residuals of Linear Models: Some Monte Carlo Results for an Irregular Spatial Lattice Structure*, Wageningen: Wageningen Agricultural University, 1991.
Florax, R. and S. Rey, *The Impacts of Misspecified Spatial Weighting Structures in Linear Regression Models*, 1992 (forthcoming).
Folmer, H., *Regional Economic Policy: Measurement of its Effect*, Dordrecht: Kluwer Academic Publishers, 1986.
Folmer, H., Autocorrelation Pre-testing in Linear Models with AR(1) Errors, in: Th.K. Dijkstra (ed.), *On Model Uncertainty and its Statistical Implications*, Berlin: Springer-Verlag, 1988, pp. 39-55.
Folmer, H. and M. Fischer, *Bootstrapping in Spatial Analysis*, Paper presented at the Symposium of the IGU Working Group on Systems Analysis and Mathematical Models, Besancon, France, 1984.
Folmer, H. and P. Nijkamp, Investment Premiums: Expensive but Hardly Effective, *Kyklos*, vol. 40, pp. 43-72, 1987.
Foster, S. and W. Gorr, Adaptive Filtering Approaches to Spatial Modeling, *Modeling and Simulation*, vol. 14, pp. 29-34, 1983.
Foster, S. and W. Gorr, Spatial Adaptive Filtering, *Modeling and Simulation*, vol. 15, pp. 745-750, 1984.
Gatrell, A., Autocorrelation in Spaces, *Environment and Planning A*, vol. 11, pp. 507-516, 1979.
Getis, A., Spatial Interaction and Spatial Autocorrelation: A Cross-Product Approach, *Environment and Planning A*, vol. 23, pp. 1269-1277, 1991.
Getis, A., Screening for Spatial Dependence in Regression Analysis, *Papers of the Regional Science Association*, vol. 69, pp. 69-81, 1990.
Gilbert, C.L., Professor Hendry's Econometric Methodology, *Oxford Bulletin of Economics and Statistics*, vol. 48, pp. 283-307, 1986.
Griffith, D., *Advanced Spatial Statistics, Special Topics in the Exploration of Quantitative Spatial Data Series*, Dordrecht: Kluwer Academic Publishers, 1988.
Hendry, D.F. and J.-F.Richard, On the Formulation of Empirical Models in Dynamic Econometrics, *Journal of Econometrics*, vol. 20, pp. 3-33, 1982.
Hepple, L., A Maximum Likelihood Model for Econometric Estimation with Spatial Series, in: I. Masser (ed.), *Theory and Practice in Regional Science, London Papers in Regional Science 6*, London: Pion, 1976, pp. 90-104.
Hordijk, L., Spatial Correlation in the Disturbances of a Linear Interregional Model, *Regional and Urban Economics*, vol. 4, pp. 117-140, 1974.
Jones, J.-P., Parameter Variation Via the Expansion Method with Tests for Autocorrelation, *Modeling and Simulation*, vol. 14, pp. 853-857, 1983.
Judge, G.G. and M.E. Bock, *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*, Amsterdam: North-Holland, 1978.
Judge, G.G., W.E. Griffiths, R. Carter Hill, H. Lütkepohl and T.-C. Lee, *The Theory and Practice of Econometrics*, New York: John Wiley and Sons, 1985.
King, M. and D. Giles, Autocorrelation Pre-Testing in the Linear Model: Estimation, Testing and Prediction, *Journal of Econometrics*, vol. 25, pp. 35-48, 1984.
Leamer, E.E., *Specification Searches, Ad Hoc Inference with Nonexperimental Data*, New York: Wiley and Sons, 1978.
Lovell, M.C., Data Mining, *Review of Economics and Statistics*, vol. 65, pp. 1-12, 1983.
Mosteller, F. and J.W. Tukey, *Data Analysis and Regression: A Second Course in Statistics*, Reading, Massachusetts, 1977.
Ord, J., Estimation Methods for Models of Spatial Interaction, *Journal of the American Statistical Association*, vol. 70, pp. 120-126, 1975.

Ripley, B., *Spatial Statistics*, New York: Wiley and Sons, 1981.
Savin, N., The Bonferroni and Scheffe Multiple Comparison Procedures, *Review of Economic Studies*, vol. 47, pp. 255-274, 1980.
Stetzer, F., Specifying Weights in Spatial Forecasting Models: The Results of Some Experiments, *Environment and Planning A*, vol. 14, pp. 571-584, 1982.
Upton, G. and B. Fingleton, *Spatial Data Analysis by Example*, New York: Wiley and Sons, 1985.