LINEAR STRUCTURAL EQUATION MODELS WITH SPATIOTEMPORAL AUTO- AND CROSS-CORRELATION.

Hendrik Folmer *

Peter Nijkamp **

Researchmemorandum 1982-20 Sept.

* Dept. of Geography, University of Groningen

** Dept. of Economics, Free University, Amsterdam

Paper presented at the 'Deutscher Geographentag', Mannheim, September, 1981.

This paper will be published in M.M. Fischer and G. Bahrenberg, Models in Geography, Bremer Beiträge zur Geographie, Heft 5, Bremen.

1. INTRODUCTION

Linear structural equation models with latent variables have recently received growing attention in spatial econometrics. They have been used in the framework of theory building and theory testing (see Droth and Fischer, 1981) and for measuring effects of regional policy (see Folmer, 1980, 1981, and Folmer and Oosterhaven, 1982). One of the assumptions underlying these models is the independence of observations. This assumption, however, is likely to be violated when spatiotemporal data is analyzed.
In the present paper some modifications of the LISREL approach will be proposed in order to cope with the problem of spatiotemporal auto- and cross-correlation[1)]. It will also be shown that LISREL models possess attractive features to deal with specific problems involved in spatiotemporal models.

The organization of this paper is as follows. In section 2 the basic features of LISREL models are summarized. The third section deals with the nature, identification and measurement of spatiotemporal auto- and cross-correlation. Estimation of spatiotemporal LISREL models will be discussed in section 4. In section 5 an empirical application will be given, in which effects of regional industrialization policy in the Netherlands will be assessed.

2. LINEAR STRUCTURAL EQUATION MODELS WITH LATENT VARIABLES

In the present chapter, the basic characteristics of linear structural models with latent variables, abbreviated as LISREL models, will be discussed. This class of LISREL models has been developed by especially Jöreskog (1973a, 1973b, 1977), Goldberger (1972), and Goldberger and Duncan (1973). Furthermore, attention will be paid to the LISREL V computer program (Jöreskog and Sörbom, 1981), which has been designed to estimate and test the current type of models. This section is based mainly on the above mentioned references.
In section 2.1., the structure of LISREL models is described. In the next section, attention is paid to the problem of identification and the way it is dealt with in LISREL V. The types of estimators used in LISREL V are described in section 2.3. These estimators are instrumental variables and non-iterative unweighted least squares, maximum likelihood, and iterative unweighted least

squares. In section 2.4., the attention is directed to the estimation of residuals, which are needed when spatiotemporal correlation is dealt with. Evaluation of estimated models and ways to improve deficient models is the theme of the final section of this chapter.
Before commencing this task, a few remarks about the nature of structural equation models have to be made. In these kinds of models the phenomena under study are described in terms of a tentative set of *cause* and *effect* variables and their relationships. Each equation in the model describes a causal relationship, so that the structural parameters represent relatively autonomous and invariant effects on the dependent variables. Consequently, the structural parameters do not, in general, coincide with regression coefficients, which usually represent merely empirical association. It is obvious that theory and prior knowledge are of great importance for the formulation of structural equation models. Since latent variables are the building stones of theory, these kinds of variables play an important role. LISREL models provide the possibility of taking them explicitly into account.

2.1. THE STRUCTURE

Let $y = (y_1, y_2, \ldots, y_p)^T$ and $x = (x_1, x_2, \ldots, x_q)^T$ be two vectors of *observable* variables. Furthermore, let $\eta = (\eta_1, \eta_2, \ldots, \eta_m)^T$ and $\xi = (\xi_1, \xi_2, \ldots, \xi_n)^T$ be vectors of *latent* variables, and $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)^T$ and $\delta = (\delta_1, \delta_2, \ldots, \delta_q)^T$ vectors of *measurement* errors of y and x, respectively.

Before specifying the relationships between these vectors of variables, the following remark is in order. It is possible to estimate intercept terms of the equations (2.1) - (2.4), describing the relationships between the variables mentioned. Such parameters may be of interest in comparisons of different, mutually exclusive, sets of regions. In the present study, however, attention will only be paid to the analysis of single samples. In such analyses, the intercept terms hardly provide any information. Therefore, we shall make the assumption here, that both the observed and the latent variables are centralized.

The relationships between the observed and latent variables are given in the latent variables measurement models (2.1) and (2.2) :

$$y = \Lambda_y \eta + \varepsilon \qquad (2.1)$$

and :

$$x = \Lambda_x \xi + \delta \qquad (2.2)$$

where Λ_y and Λ_x are $(p \times m)$ and $(q \times n)$ matrices of regression coefficients.

The structural model consists of a set of relationships among the latent variables (where the η's are the endogenous and the ξ's the exogenous variables) :

$$\eta = \tilde{B}\eta + \Gamma\xi + \zeta \tag{2.3}$$

or :

$$B\eta = \Gamma\xi + \zeta \tag{2.4}$$

where :

$\tilde{B}$ is a $m \times m$ coefficient matrix with β_{ij} representing the effect of the j-th endogenous variable on the i-th endogenous variable;

Γ is a $m \times n$ coefficient matrix with γ_{ij} representing the effect of the j-th exogenous variable on the i-th endogenous variable;

ζ is a random vector of residuals;

$B = I - \tilde{B}$, where I is the identity matrix.

The following assumptions are made. First, for reasons of simplicity, but without loss of generality, it is assumed that B is non-singular. Thus, redundant equations are assumed to have been removed. Secondly, in addition to the above mentioned assumption of centralized observed and latent variables, it is assumed that:

$$E(y) = 0 \; ; \; E(x) = 0 \; ; \; E(\eta) = 0 \; ; \; E(\xi) = 0 \quad \text{(repeated)}$$

$$E(\varepsilon) = 0 \quad E(\delta) = 0 \quad E(\zeta) = 0$$

$$E(\eta\varepsilon^T) = 0 \quad E(\xi\delta^T) = 0 \quad E(\eta\delta^T) = 0 \quad E(\xi\varepsilon^T) = 0 \quad E(\varepsilon\delta^T) = 0 \tag{2.5}$$

$$E(\zeta\xi^T) = 0 \quad E(\zeta\delta^T) = 0 \quad E(\zeta\varepsilon^T) = 0$$

The following notation is introduced. The covariance matrices of ε and δ, which need not be diagonal in LISREL, will be denoted by $\theta_\varepsilon(p \times p)$ and $\theta_\delta(q \times q)$ and the covariance matrices of ξ and ζ by $\phi(n \times n)$ and $\psi(m \times m)$.

We will conclude this section with the following remarks. First, multiple observable variables for a latent variable are preferable and often necessary in order to provide a tool for identification (see, among others, Goldberger (1972, 1977)). Secondly, one single observable variable may be an indicator of more than one latent variable. Thirdly, when combinations of observable or latent variables are omitted, several submodels of the general model (2.1) - (2.3) can be obtained. The most common of these models are first- and second-

order factor analytic models, structural equation models for directly observable models, multivariate regression models. (For details see Jöreskog and Sörbom, 1981). Fourthly, one may 'fix x' , i.e., study the conditional distribution of the y variables for given x . In that case: $\xi \equiv x$, so that $\Lambda_x = I$, the identity matrix, $\theta_\delta = 0$, and $\Phi = S_{xx}$, the sample covariance matrix of the x-variables. This type of model will be referred to as the 'fixed-x' option. If, in addition, $\Lambda_y = I$, $\theta_\varepsilon = 0$, so that $y \equiv \eta$, the structural model reads as :

$$By = \Gamma x + \zeta \tag{2.6}$$

which is the conventional simultaneous equations model.

Finally, when identification, estimation and testing of models (2.1) - (2.4) are discussed, the 'theoretical' covariance matrix Σ and the sample covariance matrix S of $z = (y^T, x^T)^T$ play essential roles. Let us first pay attention to Σ. This matrix can be expressed in terms of the eight parameter matrices Λ_y, Λ_x, B, Γ, Φ, ψ, θ_ε and θ_δ . With the specifications of (2.1) - (2.4), Σ can be written as :

$$\Sigma = \begin{bmatrix} \Lambda_y(B^{-1}\Gamma\Phi\Gamma^T(B^{-1})^T + B^{-1}\psi(B^{-1})^T\Lambda_y^T + \theta_\varepsilon & \Lambda_y B^{-1}\Gamma\Phi\Lambda_x^T \\ \Lambda_x\Phi\Gamma^T(B^{-1})^T\Lambda_y^T & \Lambda_x\Phi\Lambda_x^T + \theta_\delta \end{bmatrix} \tag{2.7}$$

On the basis of prior information (expectations, theoretical considerations, etc.), certain elements in the parameter matrices, and thus in Σ, may be regarded either as free, fixed or constrained (see inter alia Johnston, 1972). The latter means that the parameters concerned are as such unknown, but are equal to one or more other parameters. All independent, free and constrained parameters contained in the matrices Λ_x, Λ_y, B, Γ, Φ, ψ, θ_ε and θ_δ will be denoted by the vector θ . When Σ has to be expressed explicitly as a function of θ , we will write $\Sigma(\theta)$; otherwise the argument will be omitted. The sample covariance matrix is defined as :

$$S = \frac{1}{M-1} \sum_{i=1}^{M} (z_i - \bar{z})(z_i - \bar{z})^T \tag{2.8}$$

where $z_1, z_2, \ldots, z_M$ are M observations of $(y^T, x^T)^T$, and where $\bar{z} = (\bar{y}^T, \bar{x}^T)^T$ is the sample mean vector. When the observed variables have been measured on an interval or ratio scale the sample covariance matrix can be calculated in the usual way. However, when there are ordinal or nominal

variables among the observed variables special adjustments have to be made. In the case of nominal x-variables, the fixed-x option, briefly mentioned above, can be used. This is the only relevant case for the present study.

For the sake of completeness, we notice that the fixed-x option may also be used when there are ordinal variables among the x-variables. When there are ordinal variables among the y-variables or among the x-variables, which may not be considered as fixed, the LISREL program can estimate and analyze the matrices of polychoric and polyserial correlation coefficients. The first is used when all the observed variables are ordinal, and the second in case of mixed ordinal and interval data. (For further details see Olsson (1979), and Olsson et al. (1981) ; see also Barkema and Folmer (1982) for additional information).

2.2. IDENTIFICATION

The rationale behind the specification of econometric models in general, and thus also of LISREL models, is that a certain specific structure of parameters has generated the observations under consideration. The major objective of the analysis of the data gathered is to estimate the unknown parameters and to test certain restrictions on them. Estimation of LISREL models is done by fitting the theoretical covariance matrix Σ to the observed covariance matrix S of the observed variables z. The estimation procedure is based on the assumption that the distribution of the observed variables of LISREL models is adequately described by the moments of first and second order. Because the variables are assumed to be centralized (see 2.1), the distribution of z is characterized by the independent parameters in Σ. Therefore estimation of LISREL models is done by fitting Σ to S.
In order to be able to draw inferences for the vector θ from the variance-covariance matrix of the observed variables, the structure of Σ has to be such as to allow θ to be solved uniquely from Σ. Thus, the vector θ has to be determined uniquely by Σ ; in other words, the model has to be identified. With these notions it can easily be understood that if two or more different θ's yield the same Σ, the model under consideration is not identified. However, if a *single* parameter has the same value in all the θ's that generate the same Σ, the parameter is identified. If all the parameters of the model are identified, the whole model is identified.

A necessary condition for identification is that the number of distinct elements in Σ is at least as large as the number of parameters to be estimated. Let this latter number be h. There are $\frac{(p+q)(p+q+1)}{2}$ equations in h unknowns. Therefore, a necessary condition for identification is that

$$h \leq \frac{(p+q)(p+q+1)}{2} \tag{2.9}$$

A second necessary condition for identification is that each individual parameter can be separated from the other parameters. This condition is often difficult to test. Furthermore, it is not sufficient. However, the LISREL program gives hints about identification problems. It calculates an estimate of the matrix of second-order derivatives of the fitting function used to estimate the model. (The fitting functions used in the LISREL program will be described in section 2.3.) . The estimate of the matrix of second-order derivatives is obtained by substituting the value of θ , for which the first-order derivatives of the fitting function with respect to the unknown parameters are equal to zero, into the matrix of second-order derivatives.
If the log-likelihood function of the random sample is used as a fitting function, minus one times the expected value of the matrix of second-order derivatives is the *information matrix*, to be denoted as $J(\theta)$. That is :

$$J_{ij}(\theta) = -E_\theta \left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log p(z;\theta)\right] \tag{2.10}$$

where $p(z;\theta)$ is the likelihood function.

Under certain regularity conditions, which are fulfilled when the observed variables are normally distributed, the following condition holds :

$$J_{ij}(\theta) = E_\theta\left[\frac{\partial}{\partial\theta_i} \log p(z;\theta) \cdot \frac{\partial}{\partial\theta_j} \log p(z;\theta)\right] \tag{2.11}$$

In the sequel, the expected value of the matrix of second-order derivatives of the fitting function will be denoted as an 'information matrix'.
If the estimated information matrix is singular, this usually means that the model is not identified (Silvey, 1970). Furthermore, the rank of the matrix indicates which parameters are not identified (Jöreskog, 1981).
Another way of checking whether or not the model is identified is to choose a set of reasonable values for the parameters, say $\tilde{\theta}$, to let the LISREL program calculate $\Sigma(\tilde{\theta})$ and estimate the parameters of the model with $\Sigma(\tilde{\theta})$

as the input matrix. The estimated parameters different from the corresponding elements in $\tilde{\theta}$ are probably not identified.
In the case of models with latent variables, the model is not identified if the latent variables have not been assigned measurement scales. When the unit of measurement of a variable is unknown, its variance cannot be calculated without further restrictions.
The nature of this kind of non-identification can be understood most easily by recalling that a change in the unit of measurement of a latent variable combined with a corresponding adjustment of its regression coefficient will produce the same value of an observed variable. That is

$$x = \lambda \xi + \delta = \frac{\lambda}{\alpha} (\alpha\xi) + \delta \qquad (2.12)$$

where α is the rescaling factor.

The easiest way to fix the measurement scale is to set one λ-coefficient equal to 1 for each latent variable. This implies that each latent variable is measured on the scale of the corresponding observable variable with the λ-coefficient equal to 1. Another way to determine the measurement scales of the latent variables is to fix their variances. This is most easily done for the ξ variables, because Φ is a parameter matrix. The variance-covariance matrix of the η-variables is not a parameter matrix. However, the η-variables are functions of the ξ- and ζ-variables. When the variance of the ξ-variables have been fixed, the η-variables can be assigned a measurement scale by fixing the diagonal elements of the ψ matrix.
We will end this section by mentioning that when there are non-identified parameters, it is usually possible to find appropriate functions of those parameters so as to render the model identified. Further information on identification can be found in, among others, Fisher (1966), Jöreskog (1976) and Aigner and Goldberger (1977).

2.3. ESTIMATION

As mentioned above, estimation of LISREL models consists of fitting $\Sigma(\theta)$ to the sample covariance matrix S of a set of M observations on $z = (y^T, x^T)^T$ by minimizing the distance between $\Sigma(\theta)$ and S in some metric. The LISREL V program (Jöreskog and Sörbom, 1981) provides three kinds of consistent estimators :

- instrumental variables and non-iterative unweighted least squares estimators, abbreviated as 'initial estimators' ;
- maximum likelihood estimators ;
- iterative unweighted least squares estimators.

It should be noted that the term 'unweighted least squares' actually means that equal weights are used.

Before describing each of these estimators we notice that three kinds of sample matrices can be analyzed in the LISREL program. The matrix of moments about zero has to be used when intercept terms and means of the latent variables are required. When the measurement scales are very different, the correlation matrix could be analyzed for numerical expediency (for an example, see section 5). Then each variable is expressed in units of its standard deviation. In all other cases the covariance matrix should be used.

It is important to know that when different types of matrices have been analyzed, the estimates are not the same in general. This applies to all three types of estimators. Furthermore, the covariance matrix in the LISREL program of the estimators, which is only given when maximum likelihood is used, is based on the analysis of a sample covariance matrix. Therefore, when a matrix of moments about zero or a correlation matrix has been analyzed, possible tests or confidence intervals have to be interpreted cautiously.

2.3.1. Initial estimators

This estimation method provides starting values so as to speed up the iterative procedures of unweighted least squares and maximum likelihood. Furthermore, it may be used in its own right, especially in those situations where rough estimates of the parameters and other information in the output are used to improve the model under consideration.

In the description of the initial estimation procedure, *instrumental variables* will be referred to twice. The basic idea of instrumental variables is to use a linear combination of the exogenous observables in the model to estimate the unknown parameters when predetermined variables and residuals (measurement errors or disturbances) are correlated. As is well-known, least squares estimators are inconsistent in case of the present kind of correlation (see, among others, Theil, 1971).

Initial estimates of the model are obtained in successive steps. First, the initial estimates of Λ_y and Λ_x are computed by means of *instrumental variables*. This estimation procedure has been developed by Hägglund (1980). The present explanation is based on this reference.

It is assumed that the scales of the latent variables have been fixed by assigning a unit element in each column of Λ_y and Λ_x. This gives a set of m linearly independent rows in Λ_y and n linearly independent rows in Λ_x. The corresponding observable variables are called reference variables. The estimation procedures for Λ_y and Λ_x are the same. Therefore we will only describe the estimation procedures for Λ_y. For ease of notation, the subindex of Λ_y will be dropped.
In addition to the assumptions mentioned in section 2.1 it is assumed that θ_ε is *diagonal* and that $p \geq 3m$. The unknown elements of the rows of Λ are successively estimated. In each step, the y variables are partitioned into three groups : the m reference variables, the variable corresponding to the row to be estimated and the other $p - m - 1$ observables. This gives :

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} I \\ \lambda_2 \\ \Lambda_3 \end{bmatrix} \eta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \tag{2.13}$$

where y_1 $(m \times 1)$ are the reference variables, y_2 (1×1) and y_3 $((p-m-1) \times 1)$ are the other observable variables, I $(m \times m)$ is the identity matrix, λ_2 $(1 \times m)$ is the row vector to be estimated, and Λ_3 $((p-m-1) \times m)$ is the matrix of other unknown coefficients. Substituting $y_1 - \varepsilon_1$ for η into the equation for y_2 in (2.13) gives :

$$y_2 = \lambda_2 y_1 + \varepsilon_1 - \lambda_2 \varepsilon_2 \tag{2.14}$$

The vector y_3 is uncorrelated with ε_1 and ε_2, so that :

$$E(y_3, (y_2 - \lambda_2 y_1)) = 0 \tag{2.15}$$

From (2.13) it is easily derived that :

$$\lambda_2 = \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{31} \left(\Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31}\right)^{-1} \tag{2.16}$$

where $\Sigma_{..}$ denotes the population covariance matrix between the subvectors of the vector of y-variables mentioned in the subindices. Replacing the population covariance matrices $\Sigma_{..}$ by the corresponding sample covariance matrix $S_{..}$ gives the estimate $\hat{\lambda}_2$.

When Λ_y and Λ_x have been estimated the population covariance matrix of z can be written as :

$$\Sigma = \hat{\Lambda} C \hat{\Lambda}^T + \theta. \tag{2.17}$$

where $\hat{\Lambda} = \begin{bmatrix} \hat{\Lambda}_y \\ \hat{\Lambda}_x \end{bmatrix}$, C is the covariance matrix of $(\eta^T, \xi^T)^T$, and $\theta.$ is the diagonal matrix with unique variances. Let us denote by tr(.) the trace of the matrix concerned. C and θ are estimated by minimizing :

$$f = \text{tr}(S - \hat{\Lambda} C \hat{\Lambda}^T - \theta.)^2 \tag{2.18}$$

which gives unweighted least squares estimates. Because $\hat{\Lambda}$ is given (2.18) is a neat function, which can be solved analytically. From

$$-\tfrac{1}{2} \frac{\partial f}{\partial C} = \hat{\Lambda}^T (S - \hat{\Lambda} C \hat{\Lambda}^T - \theta.) \hat{\Lambda} \tag{2.19}$$

$$-\tfrac{1}{2} \frac{\partial f}{\partial \theta.} = \text{diag } (S - \hat{\Lambda} C \hat{\Lambda}^T - \theta.) \quad , \tag{2.20}$$

it follows that :

$$\hat{C} = (\hat{\Lambda}^T \hat{\Lambda})^{-1} \hat{\Lambda}^T (S - \theta.) \hat{\Lambda} (\hat{\Lambda}^T \hat{\Lambda})^{-1} \tag{2.21}$$

and :

$$\hat{\theta}. - \text{diag } (H \hat{\theta}. H) = \text{diag } (S - HSD) \tag{2.22}$$

where:

$$H = \hat{\Lambda} (\hat{\Lambda}^T \hat{\Lambda})^{-1} \hat{\Lambda}^T \tag{2.23}$$

and diag (.) denotes the diagonal matrix concerned.

From (2.22) the elements of $\hat{\theta}.$ can be derived. Substitution of $\hat{\theta}.$ into (2.21) gives $\hat{C}$.

Finally, the structural parameters of each equation are estimated by means of two-stage least squares (see, among others, Theil, 1979). That is, if the j-th structural equation reads as :

$$\eta_j = \beta^T \tilde{\eta} + \gamma^T \tilde{\xi} + \zeta_j \tag{2.24}$$

where $\tilde{\eta}$ and $\tilde{\xi}$ are the endogenous and exogenous variables appearing in the j-th equation, then the estimates $\hat{\beta}$ and $\hat{\gamma}$ of the structural parameters are obtained from :

$$\begin{bmatrix} \hat{C}_{\tilde{\eta}\xi}\hat{C}_{\xi\xi}^{-1}\hat{C}_{\tilde{\eta}\xi}^{T} & \hat{C}_{\tilde{\eta}\tilde{\xi}} \\ \hat{C}_{\tilde{\eta}\tilde{\xi}}^{T} & \hat{C}_{\tilde{\xi}\tilde{\xi}} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \hat{C}_{\tilde{\eta}\xi}\hat{C}_{\xi\xi}^{-1}\hat{C}_{\xi\eta_j} \\ \hat{C}_{\tilde{\xi}\eta_j} \end{bmatrix} \qquad (2.25)$$

In (2.25) $\hat{C}..$ is the covariance matrix of the variables mentioned in the lower indices. It is obtained from the covariance matrix estimated in the preceding step.
From the same matrix and the parameter estimates, the elements of ψ can be obtained. The diagonal elements are :

$$\text{var } (\eta_i - \hat{\beta}^T\tilde{\eta} - \hat{\gamma}\tilde{\xi}) \qquad (2.26)$$

and the off-diagonal elements are :

$$\text{cov } (\eta_i - \hat{\beta}^T\tilde{\eta} - \hat{\gamma}^T\tilde{\xi}, \eta_j - \hat{\beta}'^T\eta' - \hat{\gamma}'^T\xi') \quad i \neq j \qquad (2.27)$$

where η' and ξ' are the endogenous and exogenous variables in the j-th equation with estimated parameter vectors $\hat{\beta}'$ and $\hat{\gamma}'$.
If the two-stage least squares procedure cannot be applied (e.g., because the number of β and γ parameters to be estimated in (2.25) is larger than the number of instrumental variables) , ordinary least squares is used. The latter may be inconsistent, however.

2.3.2. Maximum-likelihood and unweighted least squares

The maximum-likelihood estimation procedure in the LISREL program is based on the minimization with respect to the unknown parameters of the non-negative function:

$$F = \tfrac{1}{2}\,[\log|\Sigma| + \text{tr}(S\Sigma^{-1}) + \log|S| - (p+q)] \qquad (2.28)$$

by means of a modification of the Fletcher-Powell algorithm. In (2.28) $|.|$ stands for the determinant and $tr\ (.)$ for the trace of the matrix concerned. When ξ, ζ, ε, δ are multinormally distributed, and thus z, then :

$$F' = -\tfrac{1}{2}M\,[(p+q)\ \log\ 2\pi\ +\ \log\ |\Sigma|\ +\ tr(S\Sigma^{-1})] \qquad (2.29)$$

is the log-likelihood function of the sample in case of independent observations.

Under the assumption of normality of z, minimization of F (which gives the same parameter estimates as maximization of F'), results in 'genuine' maximum-likelihood estimators. Under the ususal regularity conditions, which are satisfied in case of normality, the maximum likelihood estimators of θ, say $\hat{\theta}$, are asymptotically normally distributed with mean θ and covariance matrix $\frac{1}{M}\,[J(\theta)]^{-1}$, where $J(\theta)$ is defined in (2.11) with F' substituted for $\log\ p(z;\theta)$.
In particular, they are consistent and asymptotically efficient. It is obvious from (2.28) that S has to be positive definite. This condition is satisfied when there exist no exact linear relationships between any of the z variables, and if $M > p+q$. Furthermore, the starting values needed for the minimization algorithm, say θ', should be such that $\Sigma(\theta')$, is also positive definite. The initial estimates provided by the LISREL program usually satisfy this condition.
The maximum likelihood procedure also provides an estimate of the covariance or correlation matrix of the estimators, which can be used for evaluation purposes (see section 2.5).
We want to make the following remark with respect to the use of the maximum-likelihood procedure. Although the distribution of the variables mentioned above is seldom known in practice, the assumption of multivariate normality can be defended 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 (see, inter alia, Rao, 1973). Furthermore, the maximum-likelihood procedure provides 'good' estimators for a rather wide class of distributions (Dijkstra, 1981).

The unweighted least squares estimation procedure can be justified without assumptions with respect to the distribution of the variables. The following non-negative fitting function is used in the LISREL program :

$$G = \tfrac{1}{2} \text{ tr } [(S-\Sigma)^2] \tag{2.30}$$

It is minimized with respect to the unkown free and constrained parameters in an iterative procedure. The present procedure does not provide standard errors for the estimators. It is not necessary for the matrices S and Σ to be positive definite.
We will end here with the following remark. If there are several local minima for the maximum-likelihood or unweighted least squares fitting functions, the minimization procedures may converge to local minima. An indication of a local minimum is that the solution for θ is at the boundary of, or outside of, the admissible space. This is reflected by, e.g., negative variances for the measurement errors and disturbances, extreme parameters estimates, etc. Inadmissible solutions may also occur when a small sample is analyzed (see, among others, Lawley and Maxwell, 1963) or when the model fits the data badly (For an extensive description of model evaluation and model modification see section 2.5).

2.4. ESTIMATION OF THE RESIDUALS

It will be indicated in the next chapter, that the residuals of a LISREL model are of great importance in connection with spatiotemporal cross-correlation. The vector of calculated residuals, say e_m, for the m-th observation is defined as :

$$e_m = y'_m - \hat{y}_m \tag{2.31}$$

where :

- y'_m is the vector of observed values for the observable endogenous variables ;
- $\hat{y}_m$ is the vector of LISREL estimates of the observable endogenous variables.

The residuals e_m are not given by the LISREL program, but can be estimated in the following way. First the minimum variance linear estimator $\hat{\xi}$ of ξ given x is obtained by minimizing :

$$E(\xi - Ax)^T (\xi - Ax) = \text{tr}\Sigma_{\xi\xi} - 2 \text{ tr}(A\Sigma_{\xi x}) + \text{tr}(A^T A\Sigma_{xx}) \tag{2.32}$$

with respect to A. This gives :

$$A = \Sigma_{\xi x} \Sigma_{xx}^{-1} \tag{2.33}$$

From the definitions and models given in section 2.1, it follows that :

$$A = \Phi\Lambda_x^T (\Lambda_x \Phi\Lambda_x^T + \theta_\delta)^{-1} \tag{2.34}$$

and:

$$\hat{\xi} = Ax \tag{2.35}$$

Next, the following standard estimators for η can be derived (see also Goldberger et al., 1971) :

$$\hat{\eta} = B^{-1} \Gamma \hat{\xi} \tag{2.36}$$

Finally, the standard estimator for y is :

$$\hat{y} = \Lambda_y \hat{\eta} \tag{2.37}$$

In practice, the parameter matrices in (2.34), (2.36), (2.37) are unknown and are to be replaced by their LISREL estimates.

2.5. MODEL JUDGEMENT AND MODEL MODIFICATION

The purpose of model judgement is to judge how well an estimated model fits to the sample data. Two extreme forms can be distinguished. First, *genuine hypothesis testing* according to the rules of statistical decision theory (see, among others, Ferguson, 1967). Essential to this form of evaluation is the availability of a *given* hypothesis which is tested on the sample data.

The second form will be denoted here as *assessment of model fit*. It presents itself in investigations the purpose of which is to find a model that fits the data at hand. In such analyses the same data is used over and over again. In these kinds of exploratory studies, judgement statistics are needed which reflect the change in fit of the results of models successively entertained. In such cases hypothesis testing is less appropriate because the accurracy of the estimator of a data-instigated model will be overestimated to an unknown extent (see, among others, Leamer (1978), Dijkstra (1981)).

It is superfluous to mention that in practice usually a mixture of both extreme forms of model judgement occurs. The degree to which a given investigation is a mixture should be reflected in the interpretation of the judgements.

The LISREL program produces several statistics, which can be used for a model judgement. All these statistics can be used to compare the fit of the model successively estimated. Some of them can be used for testing hypotheses under appropriate conditions.
We will start with an overview of the LISREL statistics which can be used to judge a model . Next attention will be paid to the assessment of a fit and the way in which a model under consideration can be improved. Finally hypothesis testing will be described.

2.5.1. LISREL judgement statistics

The statistics provided by the LISREL program are primarily related to :

- individual parameters;
- separate equations of the measurement models and of the structural model;
- the latent variables measurement model for the endogenous and the exogenous variables jointly;
- the structural model ;
- the model as a whole.

The statistics which relate to the individual parameters are parameter estimates and, when maximum likelihood has been used, standard errors and correlations of the estimators of the individual parameters.
For the equation of each observed variable in each latent variables measurement model the *squared multiple correlation* is given. It is defined as:

$$1 - \frac{\hat{V}_{ii}}{S_{ii}} \tag{2.38}$$

where $\hat{V}_{ii}$ is the error variance and S_{ii} the observed variance of the i-th variable. This statistic is a measure of the validity and reliability of the observed variable as an indicator of the corresponding latent variables.
The squared multiple correlation for the i-th structural equation is defined as :

$$1 - \frac{\hat{\psi}_{ii}}{\hat{var}(\eta_i)} \tag{2.39}$$

It should be noted that the statistic (2.39) does not have the same interpretation as the conventional coefficient of determination R^2 . This is because η_i is a function of both exogenous and endogenous variables, the latter of which may be correlated with ζ_i . As a consequence of this correlation, (2.39) may be negative.

A statistic, which does have the same interpretation as the coefficient of determination, is the coefficient of correlation, developed by Carter and Nagar (1977).

The *coefficient of determination* for the latent variables measurement model (for the endogenous and exogenous latent variables jointly) shows how well the observed variables serve *jointly* as indicators of the latent variables. It is defined as:

$$1 - \frac{|\hat{V}|}{|S|} \tag{2.40}$$

where $\hat{V}$ is the covariance matrix of the errors of the latent variables measurement model. The coefficient of determination for all structural equations jointly is defined as :

$$1 - \frac{|\Psi|}{|\mathrm{cov}(\eta)|} \tag{2.41}$$

where $\mathrm{cov}(\eta)$ is the covariance matrix of the endogenous latent variables.

It should be noted that the present statistic does not show what proportion of the variation in the endogenous variables is accounted for by the variation in the systematic part of the model. The reason is the same as the one mentioned in relation to (2.39). The coefficient of correlation for the complete system, developed by Carter and Nagar (1977) is exactly analogous to the coefficient of determination for a single equation.

Carter and Nagar (1977) also describe tests for the coefficients of correlation for the single equations and for the complete system.

For the model as a whole several statistics are provided. First, there is the χ^2-measure which is given if maximum likelihood is used. It is defined as :

$$\tfrac{1}{2} M \left[\log |\hat{\Sigma}| + \mathrm{tr}(S\hat{\Sigma}^{-1}) - \log |S| - (p+q)\right] \tag{2.42}$$

where $\hat{\Sigma}$ is the theoretical covariance matrix calculated on the basis of $\hat{\theta}$.

The number of degrees of freedom is equal to :

$$\tfrac{1}{2}(p+q)(p+q+1) - h \tag{2.43}$$

where h is the total number of independent parameters estimated in the hypothesized model. The interpretation of the χ^2-statistic will be given in the sections 2.5.2 and 2.5.3.

Another measure for the overall fit, when maximum likelihood is used, is the *goodness of fit index* (GFM) defined as :

$$GFM = 1 - \frac{tr(\hat{\Sigma}^{-1} S - I)^2}{tr(\hat{\Sigma}^{-1} S)^2} \qquad (2.44)$$

where $\hat{\Sigma}$ is the estimated theoretical covariance matrix
This measure, adjusted for degrees of freedom (AGFM), is defined as :

$$AGFM = 1 - \frac{(p+q)\ (p+q+1)}{2h}\ (1 - GFM) \qquad (2.45)$$

Similar measures as (2.44) and (2.45) are given for unweighted least squares. Then GFM is replaced by GFU defined as :

$$GFU = 1 - \frac{tr(S - \hat{\Sigma})}{tr(S^2)} \qquad (2.46)$$

All measures (2.44) - (2.46) are expressions of the relative share of variances and covariances accounted for by the model. They usually fall between zero and one. A good fit corresponds to values close to one.

A measure of the average of the residual variances and covariances is the *root mean square residual*. It is given both when maximum likelihood and when unweighted least squares is used. It is defined as :

$$\left[2 \sum_{i=j}^{p+q} \sum_{j=1}^{i} (S_{ij} - \hat{\sigma}_{ij})^2 \ /\ (p+q)\ (p+q+1)\right]^{\frac{1}{2}} \qquad (2.47)$$

A small value of this statistic in relation to the sizes of the elements in S is an indication of a good fit.

Finally, the LISREL program gives *normalized residuals* which are approximately standard normal variates. A normalized residual is defined as :

$$\frac{S_{ij} - \hat{\sigma}_{ij}}{\left[\frac{S_{ii} S_{jj} + S_{ij}^2}{N}\right]^{\frac{1}{2}}} \qquad (2.48)$$

As a rule of thumb, a normalized residual larger than 2, is an indication of specification errors. (For further details see section 2.5.2).
The program can also give a summary of the normalized residuals jointly in the form of a so-called Q-plot. This is a plot of the normalized residuals against normal quantiles. A slope of the plotted points equal to

or smaller than 1 is an indication of a moderate or a poor fit. Non-linearities in the plotted points are an indication of specification errors or of deviations from normality.

2.5.2. Assessment of model fit

Model deficiencies may range from fundamental misspecifications, such as near non-identification, to marginal defects.
Let us first pay attention to nearly non-identified parameters. This kind of problem may arise as a consequence of the fact that a given individual parameter is difficult to be separated from the other parameters. Nearly non-identified parameters usually reveal themselves in extremely large standard errors of and high correlations between the estimators of the parameters concerned. Near non-identification can usually be solved by specifying linear equality constraints between the parameters of which the estimators are highly correlated.
Let us now turn to other kinds of specification errors. The following types can be distinguished :

a) specification errors with respect to the distribution of the variables;
b) parameters which are incorrectly fixed (usually at zero);
c) parameters which are incorrectly assumed to be different from zero and thus are incorrectly specified as free parameters;
d) specification errors with respect to the form of the model, i.e. non-linearities in models which are assumed to be linear;
e) missing variables.

In section 2.3 the following remarks have been made with respect to the distribution of the variables. First, the assumption of normality of the observed variables is of importance when 'genuine' maximum likelihood estimators are wanted. Methods for assessing multivariate normality and for robust estimators of covariances can be found in, among others, Gnanadesikan (1977).
Secondly, the maximum likelihood procedure may also be used when the distribution of the observed variables deviates form normality. However, in that case the tests to be discussed in the next section are not quite valid and the results should be interpreted cautiously. Alternatively, the jackknife [7] can be used for the purpose of hypothesis testing (see, among others, Barkema and Folmer, 1982 and Gray and Schucany , 1972). This method can also be applied when unweighted least squares, which does not require distributional assumptions and does not give standard errors, is used.

So , we may conclude that the distributional assumption does not seriously limit the LISREL approach.

The other types of specification errors mentioned above ad (b) - (e),are generally reflected in unreasonable values of one or more of the statistics mentioned in section 2.5.1.: variances, squared multiple correlations or squared multiple coefficients which are negative; correlations which are larger than one in absolute value; covariance or correlation matrices which are not positive definite.
When one or more of the statistics mentioned above have unreasonable values two problems arise:

- which type of specification error has been made;
- when the error is of the type b) or c), the misspecified parameters have to be detected, and when the error is of the type e) the missing variables have to be counterbalanced.

Both problems are highly dependent on the nature of the investigation. The more explorative the investigation, the higher the uncertainty with respect to the form of the model, the relevant variables and the status of the parameters.

Let us pay attention to parameters which have incorrectly been fixed. Incorrectly fixed parameters usually lead to inconsistent and biased estimators for all parameters.
In case of doubt about the status of a parameter, it is usually fixed according to the 'principle of parsimony' (Box and Jenkins, 1976).
When one or more judgement statistics have unreasonable values, as mentioned above, a first step to improve the model may be to relax the 'suspicious' parameter. In addition to prior theoretical or ad hoc knowledge the *modification indices* and the *normalized residuals* may be used to detect the suspicious parameters. The modification index, given for each fixed and constrained parameter, is defined as :

$$\frac{M(\text{fod})^2}{2\ \text{sod}} \qquad (2.49)$$

where fod and sod are the first- and second-order derivatives of the fitting function with respect to the fixed or constrained parameter. When maximum likelihood is used, (2.49) is equal to the expected decrease in χ^2 if the corresponding constraint is relaxed and all estimated parameters are held fixed at their estimates (Sörbom and Jöreskog, 1982).

Under these conditions the parameter with the largest modification index in absolute value will improve the model maximally.
As shown by Dijkstra (1981), the modification indices may at best give indications and they should only be applied to parameters which could be relaxed from a theoretical point of view.
For the normalized residuals, Jöreskog and Sörbom (1981) give as a rule of thumb that an absolute value larger than 2 may be an indication of a parameter that is incorrectly fixed. The indices i and j may indicate that the equations in which the i-th and j-th variables are present (either directly, or indirectly via the corresponding latent variables) contain the parameters that are incorrectly fixed.
When maximum likelihood has been used, the correct relaxation of a fixed or constrained parameter is reflected in a large drop in the χ^2 value compared to the loss of degrees of freedom. Furthermore, the other relevant statistics also show substantial improvements. On the other hand, changes in the χ^2 value which are close to the loss of degrees of freedom may have no real significance. The same applies to minor changes in the other relevant statistics.
When iterative unweighted least squares is used, the correct or incorrect relaxation of fixed parameters is also reflected in substantial, respectively minor improvements of the relevant statistics.

Let us now pay attention to incorrectly specified, free parameters. Such parameters have no influence on consistency, provided the model is identified.
The most serious consequence of incorrectly free parameters is that the estimators are not optimal.
The way of dectecting and handling incorrectly free parameters, is almost the opposite of the way of detecting and handling incorrectly fixed parameters.
Indications can be found in the values of the estimates and, when maximum likelihood has been used, in the standard errors. When the model is re-estimated with the suspect parameters fixed, the relevant statistics should not show a substantial decrease in quality.
When there are no further indications of incorrectly fixed, constrained or free parameters and when one or more judgement statistics still have unsatisfactory values, some of the relationships in the model may be non-linear or non-additive, or essential variables may be missing. The former types of specification errors may be detected by calculating the residuals (2.31) and testing them for randomness, e.g. by the turning points test (Kendall, 1973).
Many of the non-linear or non-additive relationships can be transformed into linear ones. The graph of the residuals may give hints about the kind of transformation to be used. In the case of non-linearities a logarithmic or reciprocal transformation may be helpful (see, among others, Johnston, 1972).

When relevant variables have been omitted, the estimated coefficients may be seriously biased. Furthermore, the residual variance will have an upward bias. Therefore, substantial residual variance may be an indication of omitted relevant variables. (For detailed information on the present kind of specification errors, see among others, Theil (1971), Dhrymes (1978) .)
If it is known which variables are missing but if no data are available, the use of proxy variables may be suitable (Dhrymes, 1978).

2.5.3. Hypothesis testing

When the variables η, ξ, ε, and δ are multinormally distributed, when the sample size is sufficiently large and when a covariance matrix is analyzed to investigate a given theory, the model judgement may take the form of 'genuine' hypothesis testing. It is assumed that maximum likelihood is used.

Let us first pay attention to testing the overall fit of a model.
In large samples this can be done by a likelihood-ratio test. Let :

$$H_0 : \Sigma = \Sigma(\theta)$$

and: (2.50)

$$H_a : \Sigma \text{ is any positive definite symmetric matrix.}$$

In (2.50) θ is the specific structure of the model under consideration.
The likelihood-ratio test statistic is :

$$\ell = \frac{L_0}{L_a} \tag{2.51}$$

where :

$$\ln L_0 = -\tfrac{1}{2} M \left[\ln |\hat{\Sigma}| + \mathrm{tr}(S\hat{\Sigma}^{-1})\right] \tag{2.52}$$

is the maximum of the likelihood function, given the constraints imposed by H_0 . In (2.52), $\hat{\Sigma}$ stands for the estimate of $\hat{\Sigma}$ under H_0 . The denominator in ℓ is the maximum of the likelihood function over the parameter space for all identified models. This maximum is reached when $\hat{\Sigma} = S$. Thus :

$$\ln L_a = -\tfrac{1}{2} M \left[\ln |S| + (p+q)\right] \tag{2.53}$$

So the minimum of M.F is minus 2 times the logarithm of ℓ with F given in (2.28). Under certain regularity conditions $-2 \ln \ell$ is asymptotically distributed

as a χ^2-variable, with $\{\frac{1}{2}(p+q)(p+q+1) - h\}$ degrees of freedom, where h is the total number of independent parameters estimated in the hypothesized model (see Rao, 1965).
A sequence of nested hypotheses can be tested sequentially by means of likelihood ratio statistics. The difference in values of $-2 \ln \ell$ for the models under comparison is asymptotically distributed as a χ^2-variable with degrees of freedom equal to the number of independent restrictions, which is independent of the fact whether or not the less restrictive hypothesis is true (Lehmann, 1966). So, it is possible to test whether a given model is 'worse' than a less restrictive 'bad' model. It should be noted that when testing a sequence of nested hypotheses, one should start with the less restrictive hypothesis of the sequence (Malinvaud, 1970).

The maximum likelihood procedure also allows the construction of confidence sets for individual parameters. Furthermore, the validity of the sign or specific values of each parameter can be tested. Under the prevailing conditions the standardized estimator of each parameter is asymptotically standard normally distributed.
We will conclude here by noting that when a correlation matrix is analyzed, the tests described here can be applied when the fixed-x option is used.

After this general introduction to LISREL-methods, we will now focus the attention on more specific spatial aspects in the next section.

3. NATURE AND IDENTIFICATION OF SPATIOTEMPORAL AUTO- AND CROSS-CORRELATION

Suppose a sample of observations on R regions over T periods is available. Such data is called spatiotemporal data. Three kinds of correlations between the disturbances have to be considered in relation to spatiotemporal data :

- temporal autocorrelation;
- spatial autocorrelation;
- spatial cross-correlation.

Temporal autocorrelation has been extensively described in the literature (see, among others, INSEE, 1978) ; the way how to deal with it is the LISREL approach will be described in the next section.
In case of spatial *autocorrelation*, a phenomenon in a given region, say region r, is influenced by the same phenomenon in other regions from multiple and different directions in the current or in previous periods. Furthermore, there may be a reverse influence: the phenomenon on other regions may be influenced by the phenomenon in region r in the current or in previous periods.

When two different phenomena are involved, we will speak of spatial *cross-correlation.*

Let us now turn to some measures of spatial auto- and cross-correlation. The simplest version of the spatial autocorrelation coefficient, is the generalized Moran coefficient of continguity order s of timelag l related to one variable. It is defined as :

$$M_l^s\ (y,y) = \frac{\sum_{r=1}^{R} (y_{r,t} - \bar{y}_t)\ (L^s\ y_{r,t-l} - \bar{y}_{t-l})}{\{\sum_r (y_{r,t} - \bar{y}_t)^2\}^{\frac{1}{2}}\ \{\sum_r (y_{r,t-l} - \bar{y}_{t-l})^2\}^{\frac{1}{2}}} \qquad (3.1)$$

$$s = 1,2,\ldots,S$$
$$l = 0,1,2,\ldots,T$$

where :

$y_{r,t}$ is the variable under consideration in region r at time t

and

L^s is the spatial lag operator satisfying the condition that :

$$L^s\ y_{r,t} = \sum_{i \in A_{sr}} w^s_{r,i}\ y_{i,t} \quad , \qquad (3.2)$$

in which A_{sr} is the set of all regions of contiguity order s with respect to region r and w^s_{ri} is a contiguity weight between regions r and i such that:

$$\sum_{i \in A_{sr}} w^s_{r,i} = 1 \quad , \quad \forall\ r,\ s \qquad (3.3)$$

Furthermore:

$$\bar{y}_t = \frac{1}{R} \sum_{r=1}^{R} y_{r,t} \qquad (3.4)$$

The Moran coefficient of spatial cross-correlation for two variables y and x , of contiguity order s and of time lag l is defined as :

$$M_l^s\ (y,x) = \frac{\sum_{r=1}^{R} (y_{r,t} - \bar{y}_t)\ (L^s\ x_{r,t-l} - \bar{x}_{t-l})}{\{\sum_r (y_{r,t} - \bar{y}_t)^2\}^{\frac{1}{2}}\ \{\sum_r (x_{r,t-l} - \bar{x}_{t-l})^2\}^{\frac{1}{2}}} \qquad (3.5)$$

$$s = 1,2,\ldots,S$$
$$l = 0,1,2,\ldots,T$$

More details can be found in Cliff and Ord (1973), Martin and Oeppen (1975) and Hordijk and Nijkamp (1977).

It is clear that for each variable, a Moran coefficient of spatial autocorrelation can be calculated for each time lag and for each contiguity order. The same can be done for each pair of different variables with respect to spatial cross-correlation. Thus a matrix of spatial auto- and cross-correlations can be constructed (see Martin and Oeppen , 1975, and Hordijk and Nijkamp , 1977). This matrix will be denoted by C.

Positive correlation implies $M_1^s(.,.) > 0$ and negative correlation $M_1^s(.,.) < 0$, where $M_1^s(.,.)$ refers to both $M_1^s(y,y)$ and $M_1^s(y,x)$. Absence of spatial correlation of say order s' , 1' implies : $M_1^{s'}(.,.) = 0$.

In general, the moments of $M_1^s(.,.)$ are complex relationships, which are difficult to derive, so that further statistical inferences are very hard to draw. Cliff and Ord (1973) and Haggett et al. (1973) derived the first two moments of $M_1^s(y,y)$. They showed that it is asymptotically normally distributed under the hypothesis of no spatial autocorrelation, given the assumption that the variable concerned is normally distributed.
So, the hypothesis of presence of spatial autocorrelation can be tested in a straightforward way.

The moments of $M_1^s(x,y)$ are unknown in the literature, so that this coefficient cannot be used to test for the presence of spatial cross-correlation.
Before describing a possible procedure to detect and deal with spatial cross-correlation, the following remark is in order. This kind of correlation can be caused by variables explicitly <u>included</u> in the model and/or by variables represented in the <u>disturbance term</u>. Firstly, the first kind of cross-correlation is identified by the following procedure :

- Estimate the model without specifications for spatial cross-correlation but with specifications for possible spatial and temporal autocorrelation in the way to be described in section 4.2 and calculate the residuals by means of (2.31) - (2.37).
- Tests the residuals for spatial autocorrelation by means of (3.1).
- If the hypothesis of spatially correlated residuals is rejected, spatial cross-correlation need not be considered further.
- If the hypothesis concerned is <u>not</u> rejected, the matrix of cross-correlations C is calculated.

- Spatial cross-correlation indicated by the element of matrix C, largest in absolute value, is taken into account in the way to be described in section 4.1. The element concerned is denoted as C_{ij}.
- If the coefficient of the variable representing spatial cross-correlation indicated in the preceding step, is significantly different from zero, this variable is included into the model.
- Otherwise, spatial cross-correlation indicated by the element of the matrix C, next largest in absolute value, is considered.
- This searching process continues until a coefficient significantly different from zero is found and as long as the number of relevant observable variables in the model, which may be spatially cross-correlated, is not exhausted.
- The residuals of the model extended with a variable representing spatial cross-correlation are calculated, tested for spatial autocorrelation, and so on.

The procedure described above is summarized in figure 1. In this matrix s.a.c. and s.c.c. denote spatial autocorrelation and spatial cross-correlation. Furthermore, the matrix $\tilde{C}$ has as i,j th element C_{ij}, and zeroes elsewhere. Thus, except for one entry, this matrix is fully made up by zero's.

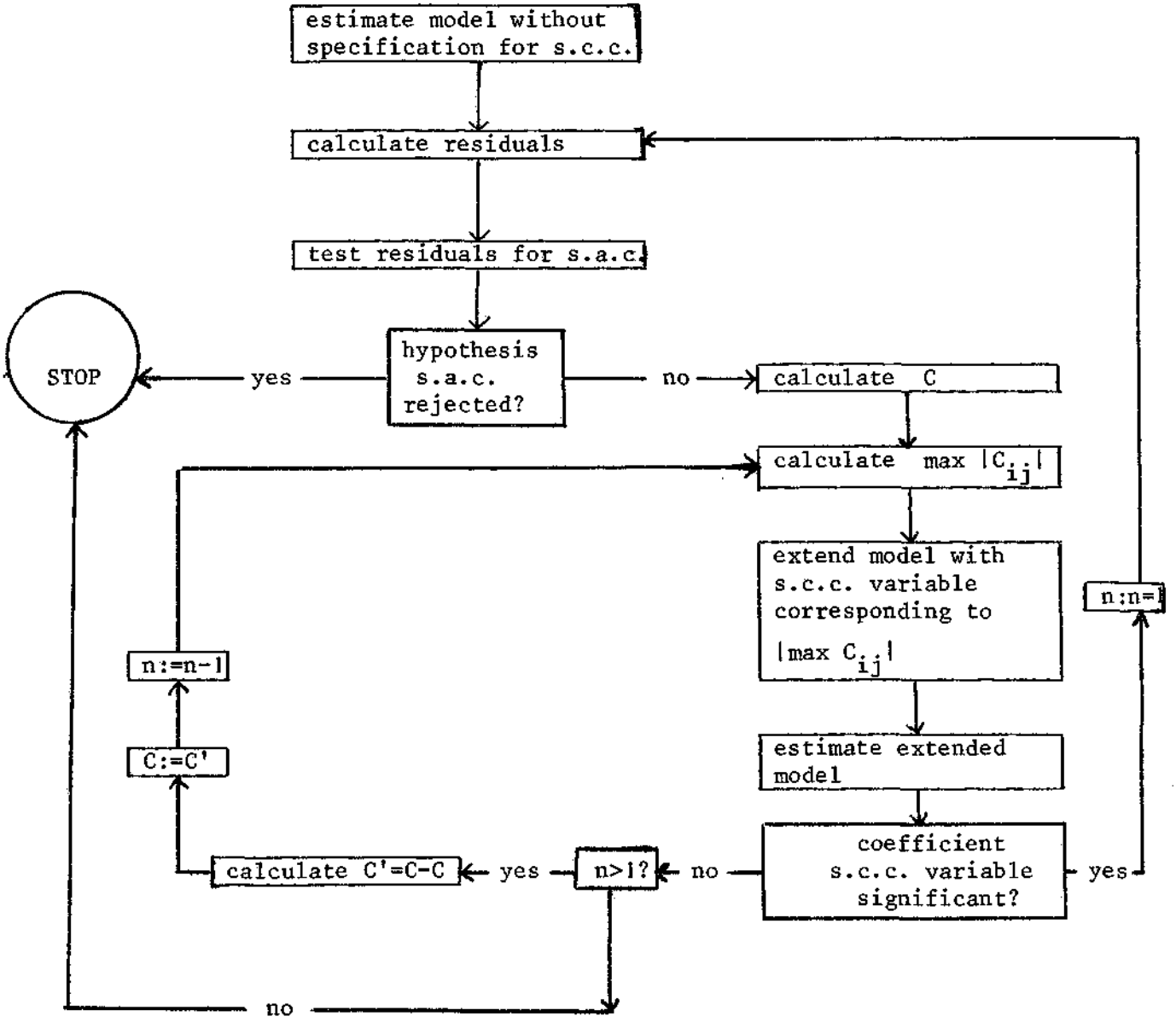

Figure 1. Schematic representation of the procedure to detect spatial cross-correlation.

If the procedure stops when the number of relevant variables is exhausted and the residuals are still spatially autocorrelated, there must be spatial auto- or cross-correlation in the variables represented by the disturbance term. It can be taken into account by procedures described by Hordijk (1974) which imply transformations of the data. The observations transformed are used to re-estimate the LISREL model.

Let us now consider (3.6) as a general specification of a model with spatial auto- and cross-correlation:

$$y_{r,t} = \sum_{l=1}^{T} \alpha_l^r \, y_{r,t-l} + \sum_{s=1}^{S} \sum_{l=0}^{L} \beta_l^s \, (L^s y_{r,t-l}) + \sum_{j=1}^{J} \sum_{l=0}^{T'_j} \gamma_{j,l}^r \, x_{r,j,t-l}$$

$$+ \sum_{j=1}^{\tilde{J}} \sum_{p=1}^{P_j} \sum_{l=0}^{\tilde{T}_j} \delta_{j,l}^p \, (L^p x_{r,j,t-l}) + \zeta_{r,t} \qquad (3.6)$$

where:

$y_{r,t}$ is the endogenous variable in region r at time t ;

$x_{r,j,t}$ is the j-th exogenous variable in region r at time t ;

α, β, γ, δ are unknown parameters ;

$\zeta_{r,t}$ represents a disturbance term.

The subscript l refers to temporal lags and the superscripts s and p to spatial lags.

Estimation of model (3.6) may be hampered by multicollinearity. One of the main consequences is that the variance of the estimators tend to increase (see among others, Johnston, 1972). Furthermore, the number of degrees of freedom may be rather low due to the lack of observations. This is a usual situation in research in geography and regional science. It will be shown in the next chapter that the LISREL approach has some attractive features to cope with these problems.

4. ESTIMATION OF SPATIO-TEMPORAL LISREL MODELS

When a test on spatial correlation does not lead to the rejection of the hypothesis of spatiotemporally independent variables, the LISREL approach described in the preceding section, can be applied in a straightforward way. Otherwise, modifications have to be introduced in order to cope with the specific data features. This will be the subject of the present section.

Suppose again that a sample of observations on R spatial units over T periods is available. The following LISREL model is considered:

$$
\left.
\begin{aligned}
B\,\eta_{r,t} &= \Gamma^{1}\,\xi_{r,t} + \Gamma^{2}_{t}\,\tilde{\xi}_{r} + u_{r,t} && \text{(a)}\\
y_{r,t} &= \Lambda_{y}\,\eta_{r,t} + \varepsilon_{r,t} && \text{(b)}\\
x_{r,t} &= \Lambda_{x}\,\xi_{r,t} + \delta_{r,t} && \text{(c)}\\
z_{r} &= \Lambda_{z}\,\tilde{\xi}_{r} + \tilde{\delta}_{r} && \text{(d)}
\end{aligned}
\right\}
\qquad
\begin{aligned}
r &= 1,2,\ldots,R\\
t &= 1,2,\ldots,T
\end{aligned}
\qquad (4.1)
$$

where :

$\xi_{r,t}$ = a vector of explanatory latent variables which may current or lagged exogenous or lagged endogenous variables;

$\tilde{\xi}_{r}$ = a vector of explanatory variables which are assumed not to vary over time ;

z_{r} = a vector of indicators corresponding to the vector of latent variables $\tilde{\xi}_{r}$ (e.g., distance from a centre) ;

$u_{r,t}$ = a vector of disturbances of the system ;

All other variables and matrices are defined as in section 2.1

4.1. LISREL SPECIFICATIONS FOR SPATIAL AUTO- AND CROSS-CORRELATION

Let us first pay attention to the way of taking spatial auto- and cross-correlation into account. Suppose there is evidence that the j-th exogenous observable variable x_{j} is spatially correlated with y_{i} for several temporal and spatial lags (see (3.6)). This kind of correlation can be dealt with by defining a new exogenous latent variable, say $\xi_{n',r,t}$, of which $L^{p}\,x_{r,t-l}$ are indicators for the various combinations of p and l concerned. Thus, instead of using a bunch of $L^{p}\,x_{r,t-l}$ variables as in equation (3.6), one latent variable representing the effects of the variable under consideration in spatial units of several orders of contiguity in several periods, is used. It it obvious that this may lead to a considerable reduction in multicollinearity.

The following remarks should be made. First, when - on the basis of prior information or data analysis - the sequence $\{L^{p}\,x_{r,t-l}\ p=1,2,\ldots,P\ ;\ l=0,2,\ldots,L\}$ is assumed to exhibit spatial or temporal transient features in its structure, more than one latent exogenous variable may be used (see Folmer and Van der Knaap, 1981, for the case of temporal transients).

Secondly, if x_j and x_h are indicators of the same latent variable and if both of them are spatially correlated with y_i, both the sequences $\{L^p x_{n,r,t-1}\}$ and $\{L^{p'} x_{j,r,t-1'}\}$ can be used as indicators of $\xi_{n',r,t}$.

An analogous approach can be used, if there is evidence of spatial auto- or cross-correlation among endogenous observable variables. In this case, a new endogenous latent variable, say $\eta_{m',r,t}$, is defined. Because it is not the intention to explain $\eta_{m',r,t}$ in terms of the other variables in the model, it is a quasi-endogenous variable. Therefore, it is set equal to its disturbance in the structural model.
The procedures described above can be applied to all observable variables for which spatial correlation holds.
A problem that may arise when R and T are small, is that the ratio of the number of observations with respect to the number of observable variables becomes equal to or smaller than one. In such cases the following solutions can be mentioned.

When a set of several indicators of a latent variable representing spatial correlation effects is available, only one indicator, which is an externally defined function of the set, could be used. Another solution is to limit the set of time lags. These procedures can be applied both to endogenous and to exogenous spatially correlated observable variables.

Finally, spatial auto- and cross-correlation in the variables represented by the disturvances can be taken into account by the methods described by Hordijk (1974). These methods can also be used to deal with spatial autocorrelation in the measurement errors of the exogenous variables.

Finally, we want to mention two other approaches to spatial correlation which are worthwhile examining in the framework of LISREL, viz. the approaches by Hordijk and Nijkamp (1978) and the one by Streitberg (1979).

4.2. LISREL SPECIFICATIONS FOR TEMPORAL AUTOCORRELATION

Let us now turn to the problem of temporal autocorrelation (see also Jöreskog (1977b) and Folmer (1981)). Let us once again consider model (4.1). The disturbance terms $u_{r,t}$ are assumed to be composed of two uncorrelated parts : a vector of individual components μ_r, which do not vary over time, and a vector of autocorrelated components ω_{rt}, which follow a first-order Markov scheme :

$$u_{r,t} = \mu_r + \omega_{r,t} \tag{4.2}$$

where :

$$\omega_{r,t} = D\,\omega_{r,t-1} + \zeta_{r,t} \tag{4.3}$$

with :

$$D = \text{diag}\,(d_1, d_2, \ldots, d_m)\ ,\ |d_i| < 1\ ,\quad i=1,2,\ldots,m. \tag{4.4}$$

$\zeta_{r,t}$ is an uncorrelated disturbance term, both within and between periods. It can be seen that $\zeta_{r,t}$ is an uncorrelated disturbance term within and between periods in the following way. Possible 'within periods' correlations would be a consequence of spatial correlation and is taken into account by means of the procedures described above. Possible 'between periods' autocorrelation is taken into account in equation (4.3).

Using Kronecker products, equation (4.1) may be combined for $t=1,2,\ldots,T$ into one single equation :

$$(I \otimes B)\,\eta_r = (I \otimes \Gamma^1)\,\xi_r + \tilde{\Gamma}^2\,\tilde{\xi}_r + u_r\ ,\quad r = 1,2,\ldots,R \tag{4.5}$$

where :

$$\left.\begin{aligned} \eta_r &= (\eta_{r,1}^T,\ \eta_{r,2}^T\ ,\ \ldots\ ,\ \eta_{r,T}^T)^T \\ \xi_r &= (\xi_{r,1}^T,\ \xi_{r,2}^T\ ,\ \ldots\ ,\ \xi_{r,T}^T)^T \\ u_r &= (u_{r,1}^T,\ u_{r,2}^T\ ,\ \ldots\ ,\ u_{r,T}^T)^T \\ \tilde{\Gamma}^2 &= ((\Gamma_1^2)^T\ ,\ (\Gamma_2^2)^T,\ \ldots\ ,\ (\Gamma_T^2)^T)^T \end{aligned}\right\} \tag{4.6}$$

I is the $T \times T$ identity matrix, while $\otimes$ denotes a Kronecker product. For simplicity of notation, the index r will be dropped. Thus, equation (4.5) reads as :

$$(I \otimes B)\,\eta = (I \otimes \Gamma^1)\,\xi + \tilde{\Gamma}^2\,\tilde{\xi} + u\ , \tag{4.7}$$

and equation (4.3) as :

$$\omega_t = D\,\omega_{t-1} + \zeta_t \tag{4.8}$$

Model (4.7) can now be formulated as a LISREL model by making the following specification. By using (4.2), the structural model is written as :

$$\begin{bmatrix} I \otimes B & J \\ 0 & A \end{bmatrix} \begin{bmatrix} \eta \\ \omega \end{bmatrix} = \begin{bmatrix} I \otimes \Gamma^1 & \tilde{\Gamma}^2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \tilde{\xi} \end{bmatrix} + \begin{bmatrix} \iota \otimes \tilde{I} & 0 \\ 0 & I \otimes \tilde{I} \end{bmatrix} \begin{bmatrix} \mu \\ \zeta \end{bmatrix} \qquad (4.9)$$

where :

μ and ζ are vectors of disturbances of dimensions m and mT respectively ;

ι is a column vector with T elements equal to one ;

$\tilde{I}$ is the $m \times m$ identity matrix ;

ω is $(\omega_1^T , \omega_2^T , \ldots , \omega_T^T)^T$, ω_i^T a m-dimensional vector;

A is a block matrix of order $mT \times mT$ of the following form :

$$A = \begin{bmatrix} I & 0 & 0 & \ldots & 0 & 0 \\ -D & I & 0 & \ldots & 0 & 0 \\ 0 & -D & I & \ldots & 0 & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & -D & I \end{bmatrix} \qquad (4.10)$$

J is a block matrix with $-D$ on the first lower diagonal and zero matrices elsewhere ; in other words, J is equal to A without the diagonal matrices I .

The measurement equation (in the LISREL-sense) for the combined $(\eta\omega)^T$ vector is specified as follows :

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_T \end{bmatrix} = \begin{bmatrix} \Lambda_y & 0 & \ldots & 0 & \bar{\Lambda} & 0 & \ldots & 0 \\ 0 & \Lambda_y & \ldots & 0 & 0 & \bar{\Lambda} & \ldots & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & & \Lambda_y & 0 & 0 & & \bar{\Lambda} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ . \\ . \\ . \\ \eta_T \\ \omega_1 \\ \omega_2 \\ . \\ . \\ \omega_T \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_T \end{bmatrix} , \qquad (4.11)$$

where $\bar{\Lambda}$ is similar in structure to Λ_y with unknown coefficients $\bar{\lambda}$.
The measurement equation for the ξ variables reads as follows :

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_T \end{bmatrix} = \begin{bmatrix} \Lambda_x & 0 & \cdots & 0 \\ 0 & \Lambda_x & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \Lambda_x \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \cdot \\ \cdot \\ \xi_T \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_T \end{bmatrix} \qquad (4.12)$$

Finally, the measurement equation of the $\tilde{\xi}$ variables has the usual form.
The correlations between ε_i and ε_j, $i \neq j$ and between δ_m and δ_n, $m \neq n$ can be expressed in θ_ε and θ_δ .

If the model described above is not identifiable, a way out may be found in restricting $\bar{\Lambda}$ or θ_ε or θ_δ . For example, θ_ε may be specified diagonal.
Ultimately the procedure described here may not be applicable.
In that case, a covariance analytical approach may be employed (see Folmer, 1981), which is another method of correcting statistically for the effects of uncontrolled variables (for time-specific features in the present case). When these corrections have been made, the usual assumptions made in LISREL, may be assumed to be satisfied.

The uncontrolled variables are generally represented by dummy variables. In case of model (4.1), this means that dummy variables are included in the vector $\xi_{r,t}$, such that :

$$\xi_{r,t} = \begin{cases} 1 & \text{for period } t \\ & \qquad r=1,2,\ldots,R \ ; \ t=1,2,\ldots,T \\ 0 & \text{for period } s \ , \ s \neq t \end{cases} \qquad (4.13)$$

The use of dummy variables has certain drawbacks (see Maddala, 1971), but these can be overcome by using 'real' information instead of dummy variables (see Folmer, 1981). In this case, a latent variable representing relevant information with respect to the various periods under consideration, has to be used (see also chapter 5).

5. AN APPLICATION : MEASUREMENT OF THE EFFECTS OF REGIONAL INDUSTRIALIZATION POLICY IN THE NETHERLANDS

In this section, the theory developed in the preceding sections will be applied to the assessment of effects of regional industrialization policy in the eleven Dutch provinces during the period 1973-1976 (for an overview of Dutch regional socio-economic policy, see among others, Oosterhaven and Folmer, 1982). The model to be estimated and tested is described in Folmer (1981). In that paper, however, absence of spatial correlation was assumed. This assumption will be abandoned here.

The following endogenous observable variables, defined for year t, are used in the model.

IO[2] : investments in buildings and transport, measured in millions of guilders;

IM[2] : investments in machinery, measured in millions of guilders ;

IPR[3] : prevailing percentage of investment premiums ;

AFD[3] : prevailing percentage of accelerated fiscal depreciation .

The latter two variables are treated as indicators for a latent variable: regional industrialization policy (RIP). It is measured on the same scale as IPR.

In addition to these real endogenous variables, two quasi-endogenous, time-invariant observable variables are used :

POP[5] : population density ;

URB[4] : degree of urbanization.

They are used as indicators for a latent variable 'social locational environment' (SLE).

The following exogenous observable variables are included into the model:

XI[4] : change of production, measured in millions of guilders ;

XIN[4] : national change of production, measured in millions of guilders;

DIS[5] : distance by road in kilometers from the 'Randstad' (the Western Metropolitan Netherlands);

IS[5] : available sites for industrial activities in hectares;

UE[6] : change in the official total unemployment percentage;

LI[4] : change in labour volume, measured in thousands of man-years

AC : first order spatial autocorrelation variable (see below).

Before presenting the results, the following remarks are in order. First, according to the theory outlined in chapter 4, the first step in dealing with spatial auto- and cross-correlation is to test for their presence. However, as mentioned in section 3, the Moran coefficient is *asymptotically* normally distributed under appropriate conditions. Cliff and Ord (1973) state that the number of observations should be larger than 50.

As in the present study only 11 provinces are involved, a different procedure will be followed. Instead of applying the 'testing procedures', spatial correlation will be assumed for variables which for theoretical reasons, and due to measurement definitions, deserve consideration. If this assumption is not correct for one or more variables, the estimated coefficient of the variables representing spatial correlation will turn out to be insignificant and the variables concerned will be deleted.

In the present study, the variables that might be spatially correlated are : change of production lagged for one year (XI(t-1)), with investments in machines (IM(t)) and with investments in buildings and transport (IO(t)) , both in the current year. The reason for this spatial correlation may be the existence of spatial input-output linkages. Furthermore, spatial correlation of contiguity order one is assumed for all periods under investigation. The latent variable representing this first order spatial correlation is constructed as the unweighted average of the change of production of those regions that have a common border with the observation unit.
Secondly, because the purpose of this model is to estimate the effects of policy, the exogenous variables are treated as fixed, (see section 2).
Thirdly, for each endogenous latent variable one λ coefficient in the measurement model has been fixed on 1 for reasons of identification (see section 2).
Fourthly, the investment variables are single indicator variables. The variances of the measurement errors of these variables were not identifiable and for that reason the covariance approach has been used. Instead of dummy variables 'real' information in the form of national investments in industry (IIN) and national change of production (XIN) have been used to represent time-specific effects.

Fifthly, a correlation matrix has been analyzed, because of the different measurement scales of the various variables. As said in section 2.3 the standard errors may not be valid in case of the analysis of a correlation matrix, so that the Z values (see below) should be interpreted cautiously.

The most important results are given below with Z-values in parentheses. The Z-value is equal to the estimated coefficient divided by its standard error under the hypothesis that the true coefficient is equal to zero. This ratio is normally distributed under the assumptions made in LISREL, so that a Z-value of 1.96 is the critical value of a two-sided test at level 5%. The measurement model for regional industrialization policy (RIP) reads as follows:

$$IPR(t) = RIP(t) + \varepsilon_1(t) \qquad (5.1)$$

$$AFD(t) = \underset{(4.6)}{0.63}\ RIP(t) + \varepsilon_2(t) \qquad (5.2)$$

The measurement model for the quasi-endogenous latent variable social locational environment (SLE) is :

$$POP = SLE + \varepsilon_3(t) \qquad (5.3)$$

$$URB = \underset{(26.3)}{1.1}\ SLE + \varepsilon_4(t) \qquad (5.4)$$

From the estimated variance-covariance matrix of the measurement errors, $\hat{\theta}_\varepsilon$, which will not be presented here, it follows that the variance of the measurement error of AFD(t) is rather large. The variances of all other measurement errors are small.

The structural equations are :

$$IO(t) = \underset{(2.8)}{0.2}\ SLE + \underset{(3.8)}{0.22}\ XI(t-1) + \underset{(1.4)}{0.1}\ DIS + \underset{(0.6)}{0.1}\ IS + \underset{(16.3)}{0.9}\ IO(t-1) - \underset{(0.7)}{0.1}\ IIN(t) +$$

$$+ \underset{(1.1)}{0.1}\ AC(t-1) + \zeta_1(t) \qquad R^2 = 0.93 \qquad (5.5)$$

$$IM(t) = \underset{(3.1)}{0.1}\ IX(t-1) + \underset{(2.0)}{0.1}\ XI(t) + \underset{(4.2)}{0.2}\ AFD(t-1) + \underset{(5.5)}{0.2}\ LI(t-1) + \underset{(17.8)}{0.8}\ IM(t-1) +$$

$$- \underset{(1.0)}{0.1}\ IIN(t) - \underset{(2.9)}{0.1}\ AC(t-1) + \zeta_2(t) \qquad R^2 = 0.95 \qquad (5.6)$$

$$RIP(t) = \underset{(1.0)}{0.1}\ UE(t) + \underset{(19.0)}{0.8}\ IPR(t-1) + \underset{(4.1)}{0.1}\ AFD(t-1) - \underset{(2.9)}{0.2}\ XIN(t) + \zeta_3(t)$$

$$R^2 = 0.92 \qquad (5.7)$$

The χ^2 of this model is 224.3 with 70 degrees of freedom. The results are quite similar to the results described in Folmer (1981) to which the reader is referred for further details. It should be noted however, that the "R^2's" , defined as $1 - \hat{\psi}_{ii}$, in the present model are higher, as might be expected. As to the spatial cross-correlation variable, its coefficient in the investment in machines equation is highly significant and it could not be deteled here. In the investment in buildings and transport equation its importance is less clear.
Summarising, we may conclude that the inclusion of the first-order autocorrelation variable means an improvement of the model.

NOTES

1. In the literature both spatial auto- and cross-correlation are often denoted by the term 'spatial autocorrelation'. In this paper, it is more convenient to use two terms instead of one.

2. Central Bureau of Statistics : Statistiek van de investeringen in vaste activa in de nijverheid, 1973-1976. The Hague.

3. Staatscourant. For provinces which only partly benefit from these regulation, the data is derived by multiplying the prevailing percentage with the population shares of the areas concerned. (for further details see Folmer and Oosterhaven, 1982).

4. Central Bureau of Statistics : Regionale Economische Jaarcijfers, 1971 and 1973-1976. The Hague. No data was available for 1972; it has been calculated as the average of 1971 and 1973.

5. Central Planning Bureau: Centraal Economisch Plan 1978. The Hague.

6. Central Bureau of Statistics : Sociale Maandstatistiek 1972-1976. The Hague.

7. The jackknife procedure may also be an alternative way of hypothesis-testing when a correlation matrix or a matrix of moments about zero is analyzed.

REFERENCES

AIGNER, D.J., A.S. Goldberger (1977): Latent Variables in Socio-economic Models. Amsterdam: North-Holland.

BARKEMA, H., H. Folmer (1982). Analysis of Discrete Variables by means of Linear Structural Equation Models. Paper to be presented at the European Colloquium on Theoretical and Quantitative Geography. Augsburg, September 1982. (Mimeographed by Department of Geography, University of Groningen).

BOX, G.E.P., G.M. Jenkins (1976). Time Series Analysis, San Fransisco : Holden Day.

CARTER, R.A.L., A.L. Nagar (1977. Coefficients of Correlation for Simultaneous Equation Systems. Journal of Econometrics 6, 39-50.

CLIFF, A.D., J.K. Ord (1973): Spatial Autocorrelation. London: Pion.

DHRYMES, Ph. J. : Introductory Econometrics. New York : Springer(1978).

DROTH, W., M.F. Fischer (1981) : Zur Theoriebildung und Theorietestung. Eine Diskussion von Grundlagenproblemen am Beispiel der Sozialraumanalyse. In OSTHEIDER, M., STEINER, D. (eds.): Theorie und Quantitative Methodik in der Geographie, Zürcher Geographische Schriften. Heft 1. Geographisches Institüt Eidgenössische Technische Hochschule Zürich.

DIJKSTRA, T.K. (1981) : Latent Variables in Linear Stochastic Models. Ph. D. University of Groningen.

FISCHER, F.M. (1970). A Correspondence Principle for Simul taneous Equation Models. Econometrica, 38, 73-92.

FISCHER, F.M. (1966) : The Identification Problem in Econometrics. New York: McGraw-Hill.

FOLMER, H. (1980) Measurement of the Effects of Regional Policy Instruments. Environment and Planning A 12, 1191-1202.

FOLMER, H. (1981) Measurement of the Effects of Regional Policy Instruments by means of Linear Structural Equation Models and Panel Data, Environment and Planning A 13, 1435-1448.

FOLMER, H., J. Oosterhaven (1982) : Measurement of the Effects of Dutch Regional Socio-economic Policy . In KUKLINSKI, A., J.G. LAMBOOY (eds.): Regional Planning and Regional Policies, Experiences and Prospects : A Polish-Dutch Seminar. The Hague: Mouton. (forthcoming).

FOLMER, H., B. van der Knaap (1981). A Linear Structural Equation Approach to Cross-sectional Models with Lagged Variables. Environment and Planning A 13, 1529-1537.

GNANADESIKAN, R. (1977): Methods for Statistical Data Analysis of Multivariate Observations. New York : Wiley.

GOLDBERGER, A.S., A.L.Nagar, M.S. Odeh (1961). The Covariance Matrices of Reduced-form Coefficients and of Forecasts for a Structural Econometric Model. Econometrica 29, 556-565.

GOLDBERGER, A.S. (1972). Structural Equation Models in the Social Sciences. Econometrica 40, 979-1002.

GOLDBERGER, A.S., O.P. Duncan (eds.) (1977): Structural Equation Models in the Social Sciences. New York : Seminar Press.

GRAY, H.L., W.R. Schucany (1972) : The Generalized Jackknife Statistic. New York: Dekker.

HAGGET, P., A.D. Cliff, A. Frey (1977): Locational Methods. London: Arnold.

HÄGGLUND, G. (1980): Factor Analysis by Instrumental Variables Methods : A Comparison of Three Estimation Procedures. Research Report 80-2. Department of Statistics. University of Uppsala, Sweden.

HORDIJK, L., (1974): Spatial correlation in the disturbances of a linear inter-regional model. Regional and Urban Economics 4, 117-140.

HORDIJK, L., P. Nijkamp (1977): Dynamic Models of Spatial Autocorrelation. Environment and Planning A 9, 505-519.

HORDIJK, L., P. Nijkamp (1978): Estimation of Spatio-temporal Model. In KARLQVIST, A. and others (eds.): Spatial Interaction Theory and Planning Models, 177-200. Amsterdam: North-Holland.

INSEE, (1978) : Annales de l'INSEE, issue 30-31 (April-September) (Proceedings of the Conference on Econometrics of Panel Data, August 22-24, 1977): Paris.

JOHNSTON, J. (1972): Econometric Methods. New York: McGraw-Hill.

JÖRESKOG, K.G. (1973a) : A General Method for Estimating a Linear Structural Equation System. In: GOLDBERGER, A.S., and O.P. Duncan (eds.): Structural Equation Models in the Social Sciences. New York: Seminar Press, 85-112.

JÖRESKOG, K.G. (1973b): Analysis of Covariance Structures. In: KRISHNAIAH (ed.): Multivariate Analysis-III. New York: Academic Press, 263-285.

JÖRESKOG, K.G. (1977): Structural Equation Models in the Social Sciences: Specification, Estimation and Testing. In : KRISHNAIAH (ed.): Application of Statistics, Amsterdam: North-Holland, 265-287.

JÖRESKOG, K.G. (1977b): An Econometric Model for Multivariate Data in Annales de l'INSEE 30-31 (April-September), Paris, 355-367.

JÖRESKOG, K.G. (1981): Basic Issues in the Application of LISREL. Data 1, 1-6.

JÖRESKOG, K.G., D. Sörbom (1981) : LISREL V. Analysis of Linear Structural Relationships by Maximum Likelihood and Least Squares Methods. Mimeographed by Department of Statistics, University of Uppsala, Sweden.

LAWLEY, D.N., A.E. Maxwell (1963): Factor Analysis as a Statistical Method. London : Butterworth.

LEAMER, E.E. (1978): Specification Searchers. Ad hoc Inference with Non-experimental Data. New York : Wiley.

LEHMAN, E.L. (1966): Testing Statistical Hypotheses. New York : Wiley.

MADDALA, G.S. (1971): The Use of Variance Component Models in Pooling Cross Section and Time Series Data. Econometrica 39, 341-358.

MALINVAUD, E. (1970): Statistical Methods of Econometrics. Amsterdam: North-Holland.

MARTIN, R.L., J.E. Oeppen (1975). The Identification of Regional Forecasting Models Using Space-time Correlation Functions. Trans. Inst. Br. Geographers 66, 95-118.

OLSSON, U. (1979). Maximum Likelihood Estimation of the Polychoric Correlation Coefficient. Psychometrika, 44, 443-460.

OLSSON, U., F. Drasgow, N.J. Dorans (1981): The Polyserial Correlation Coefficient. Mimeographed by Department of Statistics, University of Uppsala, Sweden.

OOSTERHAVEN, J., H. Folmer (1982): Overview and Review of Dutch Regional Socio-economic Policy. In : KUKLINSKI, A., J.G. Lambooy (eds.): Regional Planning and Regional Policies, Experiences and Prospects : A Polish-Dutch Seminar. The Hague: Mouton Publishers. (forthcoming).

RAO, R. (1965): Linear Statistical Inference and its Applications. New York: John Wiley.

SILVEY, S.D. (1970): Statistical Inference. Harmondsworth, Middlx: Penguin Books.

STREITBERG, B. (1979): Multivariate Models of Dependent Spatial Data. In KETELLAPPER, R.M., C.P. BARTELS (eds.): Exploratory Analysis of Spatial Data, 139-178. Boston: Martinus Nijhoff,139-179.

THEIL, H. (1971): Principles of Econometrics. New York: Wiley.

Serie Research Memoranda:

1979-1	A.A. Schreuder-Sunderman, Frans Blommaert en Hein Schreuder,	Werknemers en Soc. Jaarverslag,
1979-2	H. **Schreuder,**	**De maatschappelijke verantwoordelijkheid van Ondernemingen.**
1979-3	**P. Nijkamp en P. Rietveld,**	**Multilevel Multi-objective Models in a Multiregional System.**
1979-4	**J. Arntzen, G. Bornmalm-JardelSw and P. Nijkamp,**	**Duality, Segmentation and Dynamics on a Regional Labour Market.**
1979-5	**P. Nijkamp, A. Soffer**	**Soft Multicriteria Decision Models for Urban Renewal Plans.**
1979-6	**drs. A.J. Mathot**	**A Model of choosing a car with or without a credit.**
1979-7	H. **Blommestein, P.** Nijkamp **en W. van Veenendaal,**	**Shopping Perceptions and preferences: A multi-dimensional Attractivenerss Analysis of Consumer and Entrepreneurial Attitudes.**
1979-8	H.J.Blommestein and F.C. Palm	The Aggregate Demand for Money in the Netherlands - a new look at a study of the Bank of the Netherlands.
1980-1 Jan.	P. Nijkamp and H. Voogd	New Multicriteria Methods for Physical Planning by Means of Multidimensional Scaling Techniques.
1980-2	**Hidde P. Smit**	Medium- and Long-Term Models for the Escap-Region - A review of existing models and a proposal for a new model system -
1980-3 april	P.v. Dijck en H. Verbruggen	Froductive Employment in Developing Countries' Exporting Industries
1980-4	P. Nijkamp en L. Hordijk	Integrated Approaches to Regional Development Models; A survey of some Western European Models
1980-5	P. Nijkamp	Soft Econometrie Models. An analysis of Regional Income Determinants
1980-6	P. Nijkamp en F. van Dijk	Analysis of Conflicts in Dynamical Environmental Systems via Catastrophe Theory.
1980-7 juni	E. Vogelvang	A short term econometrie model for the consumer demand of roasted coffee in the Netherlands
1980-8	N. van Hulst	De effectiviteit van de geleide loonpolitiek in Nederland.
1980-9 Oct. 1980	P. Nijkamp	A survey of Dutch integrated Energy-Environmental-Economic Policy Models
1980-10 Oct. 1980	P. Nijkamp	Perspectives for Urban analyses and policies
1980-11 Oct. 1980	P. Nijkamp	New developments in multidimensional geographical data and policy analysis
1980-12	F.C. Palm, E. Vogelvang en D.A. Kodde	Efficiënt Estimation of the Geometrie Distributed Lag Model; Some Monte Carlo Results on Small Sample Properties