CAUSALITY ANALYSIS IN SOFT SPATIAL ECONOMETRIC MODELS

Hans Blommestein* and Peter Nijkamp**

1. *Introduction to Soft Data Methods*

Traditionally, urban and regional models are largely based on the following (non-exhaustive) set of assumptions:

(i) a well-defined set of variables can be measured by means of a cardinal metric;

(ii) the complex relationships among these variables can be quantified by means of an operational economic model that describes the relevant causal impacts;

(iii) the technical, institutional, social and economic conditions of the system concerned are precisely known and can be specified in an operational way;

(iv) the spatial spillover effects can be estimated precisely via a spatial distributional model;

(v) wherever uncertainty exists concerning the state of the system, the probability distribution of stochastic elements is assumed to be known;

(vi) for any dynamic model, the time trajectory of all variables of the spatial system can be computed precisely.

A natural conclusion from this might be that full and precise information is a basic ingredient for traditional spatial econometric model building. In recent years, however, several authors have argued that some phenomena can hardly be quantified by means of a cardinal metric system (see, among others, Nijkamp, 1979, 1980a; Blommestein and Van Deth, 1981). Others have indicated that cardinal information may imply a serious bias, so that limited reliability can be claimed (see, among others, Adelman and Morris, 1974). The level of measurement therefore deserves a close examination. In general, the following measurement scales may be distinguished: *quantitative* (or *cardinal*) information subdivided into *ratio* and *interval* data, and *qualitative* (or *categorical*) information subdivided into *ordinal* and *nominal* data. In this paper, *soft* information may be interpreted as either qualitative or fuzzy data (fuzziness implies that the boundaries of the measurement space of variables are not exactly known, but must be approximated by means of *cardinal membership grades* (see Zadeh et al., 1975).

Fuzzy set problems fall beyond the scope of this paper, where the attention will be focussed on qualitative data. Qualitative data have received much attention in psychology, sociology and biology, but less attention in economics (although

* Twente University of Technology, The Netherlands.
** Free University, The Netherlands.

the interest in disaggregate choice models has stimulated the use of soft data techniques). The following qualitative data methods may be identified:

(i) ordinal correlation analysis via rank correlation coefficients;
(ii) dummy variable techniques;
(iii) path models via proxy variables (Blalock, 1964): useful methods are Lisrel (see Folmer, 1980, and Jöreskog, 1977) and Partial Least Squares (see Wold, 1975);
(iv) multidimensional scaling methods transforming ordinal data into a smaller set of cardinal variables;
(v) categorical data analysis based on disaggregate information (see Wrigley, 1980);
(vi) multinomial logit and probit analysis for discrete data on qualitative attributes or objects (see Theil, 1971, Domencich and McFadden, 1975, Van Lierop and Nijkamp, 1980);
(vii) ordinal regression analysis based on various modes of stochastic orderings of an ordinality structure (see McCullagh, 1980);
(viii) contingency table analysis (see Grizzle et al., 1969, and Lehnen and Koch, 1974);
(ix) pairwise logit transformation via a regime method (see Nijkamp and Rietveld, 1980).

All these methods serve to draw quantitative inferences from qualitative data. The specific question to be dealt with in this paper is how such qualitative data can be analysed in the framework of causal spatial economic models. The next section will be devoted to a brief introduction to causality analysis.

2. *Causality Analysis*

A number of urban and regional economic models take for granted the existence of well-defined relationships for the quantitative impacts of causal variables upon effect variables. Causality analysis has attracted much attention in the social sciences, particularly in economics (see, among others, Simon, 1953, 1954, 1957; Lazarsfeld, 1954; Wold, 1954; Kendall, 1955; Fox et al., 1966; Harvey, 1969; Rietveld, 1981; Blommestein, 1981a; and Nijkamp and Rietveld, 1982).

Causality can be understood in two ways: (1) as *relational* causality, in order to test by means of statistical and econometric techniques the existence of cause-effect links; (2) as *structural* causality, which characterizes the overall structure of an economic system by means of a coefficient matrix of transition. Causality analysis is also extremely important in the analysis of regional and urban models, especially as far as the presence of a top-down or bottom-up relationship is concerned. In complex regional and urban models, a closer examination of causality structure may illuminate the hierarchical impacts of cause variables upon effect variables.

In the present paper, causality will be examined in its structural sense; it is a property to be studied within the specific framework of a model or a system. Hence, it does not necessarily refer to real-world impact patterns (cf. Simon, 1957). In a spatial setting, causality analysis is further complicated by the existence of spatial spillover effects leading to spatial cross-correlation problems (see among others, Cliff and Ord, 1973; Ord, 1975; Hepple, 1976; Hordijk, 1979; Nijkamp, 1979; and Blommestein, 1981b). A model can formally be represented by means of N structural relationships, as follows:

$$h\ (y,\ x) = 0 \tag{1}$$

where y is an $(NX1)$ vector with endogenous variables, x a $(KX1)$ vector with exogenous variables, and h an implicit vector function. The specification of function h depends on a set of hypotheses H. System (1) represents a particular structure of the pair of sets $\{z,\ H\}$; Gilli (1980) notes that the *causal structure* of model (1) is already given if the set of hypotheses H defines the partition $z = y \cup x$, as well as a set of binary relationships of the form $h_i R z_j$, $h_i \epsilon H = \{h_1, h_2, \ldots, h_I\}$ and $z_j \epsilon z = \{z_1, z_2, \ldots, z_N, z_{N+1}, \ldots, z_{N+K}\}$; the operator R stands for: "the i[th] relation contains variable z_j." Alternatively, the causal structure can be defined by zero entries in the matrix of first-order derivatives (cf. Rietveld, 1981), i.e.,

$$D = \left[\frac{\partial h}{\partial y'} \;\middle|\; \frac{\partial h}{\partial x'} \right] \tag{2}$$

Further insight into the causality structures implied by (2) can be gained by examining the *sign* of the non-zero elements. The information obtained in this way constitutes a *calculus of qualitative relations* (Samuelson, 1947). It should be noted that the zero and non-zero entries of (2) can be derived independently from a specific quantification of relationship h. Despite the limited information regarding the specific quantification of the set of hypotheses H, substantial prior information must be available for defining partitions like $z = y \cup x$. If the latter information is not available, it is necessary to conduct *inter alia* more empirical research or to gather additional functional information.

Causality analysis can be carried out at different levels of measurement. Consider, for example, the following simple structural model with three endogenous variables, namely income (Y), investment (I) and consumption (C), and three exogenous variables, namely government expenditures (G), taxes (T) and the interest rate (d):

$$\begin{aligned} h_1\ (C,\ Y,\ T,\ d) &= 0 \\ h_2\ (W,\ C,\ I,\ G) &= 0 \\ h_3\ (I,\ Y,\ d) &= 0 \end{aligned} \tag{3}$$

where these equations represent the consumption relationship, the income identity relationship, and the investment relationship, respectively. Setting the non-zero elements of the matrix with the first-order derivatives of each endogenous variable of system (3) with respect to a cause variable equal to one, yields a *causal* structure in the form of the adjacency matrix A for the successive variables W, I, C, G, T and d:

$$A = \begin{bmatrix} 1 & 1 & 0 & | & 0 & 1 & 1 \\ 1 & 1 & 1 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

The *qualitative* structure of model (3) can be expressed in the form of a sign matrix S, with elements $+$, $-$ or 0. If each endogenous variable (W, I and C) is expressed explicitly in terms of the successive remaining variables, S will have the following structure:

$$S = \begin{bmatrix} 0 & + & 0 & | & 0 & - & - \\ + & 0 & + & | & + & 0 & 0 \\ 0 & + & 0 & | & 0 & 0 & - \end{bmatrix} \tag{5}$$

In (2) and (3), it was necessary to assume that all variables were measured on a metric scale. The alternative assumption, that some or all variables are measured on a non-metric scale, requires information about the direction of the impacts before the causality analysis can be applied. For example, the first relationship in (3) can be expressed as follows:

$$h_1 \{f_C^o (C), f_W^o (W), f_T^o (T), f_d^o (d)\} = 0 \tag{6}$$

in which $f_x^o (x)$ is an arbitrary order-preserving function of an ordinal variable x with properties ($r_i \varepsilon \{1, \ldots, R\}$ indicates the i[th] rank order of variable x):

$$\begin{aligned} f_x^o (r_i) &> f_x^o (r_{i'}), \text{ if } r_i > r_{i'} \\ f_x^o (r_i) &= f_x^o (r_{i'}), \text{ if } r_i = r_{i'} \\ f_x^o (r_1 = 1) = 1&;\ f_x^o (r_i = R) = R \end{aligned} \tag{7}$$

One can go further than the calculus of qualitative *relations* by calculating the weights or parameters which measure the relative strength of the relationship(s) between the variables. It is then necessary to construct causal structural models dealing with *non-metric variables*. This leads us into the area of soft econometrics, alluded to in the previous section.

An analogous approach can be followed for causality analysis in a spatio-temporal context. It has been stressed by several authors (Bennett and Chorley, 1978; Harvey, 1969; and Blalock, 1964) that the notion of time is crucial for analyzing, understanding and interpreting causal orderings. Although timebased systems share the conditions of asymmetry and transitivity, the spatial simultaneity phenomenon in purely space-based system prevents the fulfilment of these conditions (see Bennett, 1979). In general, the notion of causality may be meaningful in both temporal and spatial systems, but the operationalisation of this concept in spatial systems is more complicated (see Basman, 1963; and Blommestein and Nijkamp, 1981).

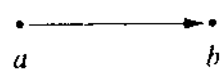
After the discussion of soft (spatial) data methods and a formal framework for causality analysis in economic models, we can examine the possibility of studying causality patterns of soft (spatial) economic models by means of techniques such as graph theory. This will be the subject of the next section.

3. *Graph Theory and Causality Analysis*

The complex relationships in a spatial system can be represented in many forms. Causal models (including non-metric models) are being increasingly analyzed by means of path and graph analyses. Graph theory provides an attractive multilevel language for dealing with complex problems (Riguet, 1980).

To handle soft information, the analysis of complex causality structures is fraught with statistical problems. A meaningful way of analyzing a causality structure based on soft information is provided by graph theory. Graph theory is a mathematical tool for reducing the structural and functional complexity in a system to a set of systematic linkages that allow further mathematical treatment (see, among others, Behzad and Chartrand, 1971; Marshall, 1971; Christofides, 1975; Andrasfai, 1977; and Beineke and Wilson, 1978).

For the analysis of causality relationships, directed graphs (di-graphs) are important. A di-graph is a finite non-empty set of vertices V together with a set E (distinct from V) of ordered pairs of individual elements of V. The elements of E are called arcs. For instance (a,b) is an arc signifying that there is a relationship from a to b, or:



A causality relationship between a and b presupposes the existence of an arc. The *degree of causality* can be measured via the degree of the associated vertex, where the degree is related to the (unweighted) number of oriented graphs. The algebraic representation of such a causality analysis may proceed via an oriented adjacency matrix. In this way, graphs may act as a structural model with endogenous variables $y \in R^N$ and exogenous variables x:

$$h(y, x) = 0 \tag{8}$$

Such a representation is not necessarily a fully specified mathematical model, since the sets of variables are not measured on a cardinal metric. A causal structure of (8) presupposes a set h of cause-effect relationships which partitions a set of variables z into $y \cup x$. A set of relationships can then be represented by a *bi-partite* graph G (see Harary et al., 1965, and Gilli, 1980):

$$G = (h, z, q), \tag{9}$$

where $q = \{(h_i, x_j) \mid h_i R x_j\}$ is a set of edges. Clearly, the information contained in (9) can be transformed into a related adjacency matrix. Next, a causal structure for a set of relationships can be represented by means of an *oriented* graph G_o:

$$G_o = (h, z, s) \tag{10}$$

where s is a set of arcs.

In addition, one may also distinguish a *directed* graph, G_d, which is based on a well-defined contraction of the vertices.

On the basis of the above mentioned notions of graphs, various types of causality may be defined (see Gilli, 1980):

— *direct* causality:

$$z_j D z_{j'} \leftrightarrow G_d = \{\text{arc } (z_j, z_{j'})\}, \tag{11}$$

where D represents a relationship based on a directed graph.

— *immediate* causality:

$$z_j C z_{j'} \leftrightarrow G_d = \{\text{path } (z_j, z_{j'})\}, \tag{12}$$

where C defines a quasi-ordering with the following properties:

$z_j C z_j$ (reflexivity), $\forall j$

$(z_j C z_{j'}) \wedge (z_{j'} C z_{j''}) \rightarrow z_j C z_{j''}$ (transitivity)

— *mutual* causality (for endogenous variables):

$$y_n E y_{n'} \leftrightarrow \{(y_n C y_{n'}) \wedge (y_{n'} C y_n)\}, \tag{13}$$

where E is a reflexive, transitive and symmetric relationship. It should be noted, however, that a symmetric causality structure may be questionable, as this may be due to the neglect or improper treatment of the time dimension. In such cases, a recursive system may be more appropriate (see Wold, 1954).

The major advantage of graph theory is that it does not presuppose a fully specified quantitative model; so it is a useful tool for soft models. Some simple topological measures of network or graph structures will now be discussed, together with a possible causal interpretation. These measures are based on the gross characteristics of networks (see Garrison and Marble, 1962; Kansky, 1963).

The *cyclomatic* number is defined as:

$$\mu = E - V + G_s \tag{14}$$

where E is the number of edges (links) in the network, V the number of vertices (nodes), and G_s the number of subgraphs.

The cyclomatic number μ defines the number of fundamental circuits in a network. A pairwise comparison of networks with the same number of variables (vertices) may reveal approximately the differences in the causal (recursive/interdependent) structure of networks.

The *alpha-* (or redundancy) index gives additional information about the connectivity of networks. It is defined as follows:

$$\text{planar graphs}: \alpha = \frac{\mu}{2V - 5} \tag{15}$$

$$\text{non-planar graphs}: \alpha = \frac{2\mu}{(V - 1)(V - 2)}$$

Since the α-index is defined as the ratio between the observed number of circuits and the maximum number of circuits, its value provides relatively sensitive information about the causal form of networks, both in the case of a single network and pairwise comparisons of networks.

The *beta-index*, defined as

$$\beta = E/V \tag{16}$$

can be considered as a simple measure of the complexity of causal networks. The β-index differentiates between simple topological structures (with low β-values) and complicated structures (with high β-values).

The *gamma-index* is calculated by dividing the number of edges by the maximum number of edges, i.e.:

$$\text{planar graphs}: \gamma = \frac{E}{3(V - 2)} \qquad \text{non-planar graphs}: \gamma = \frac{E}{V(V - 1)/2} \tag{17}$$

Like the α-index, γ provides information about the connectivity of the network.

Finally, it is possible that *different* networks can yield the *same* values for α-, β-, γ- and μ-indices.

Therefore, it might be of interest to compare 'restricted' networks by considering only a limited number of variables, for example; by deleting identities; by differentiating between directed graphs (G_d), homogeneous graphs (G_h) and p-graphs (G_p); by focussing on well-defined aspects of the causal structure of models such as the relationship(s) between national and regional variables, final demand and supply of production factors (see Rietveld, 1981).

The abovementioned causality analysis can be adapted to a spatial system by making a distinction between the *intra*-causality and *inter*-causality of a regional system. Suppose a spatial system is composed of R regions, each region being described by means of a set of structural relations reflecting interregional linkages. The intra-causality structure of a region refers to the causality of the variables and relationships within the region at hand, whereas the inter-causality refers to the causality of the whole system (including interregional causal linkages). A poorly interwoven spatial system will have a much lower causality degree than a single region. In this way, causality analysis for modules of a spatial system can be used to obtain some insight into the connectivity of the system.

4. *Empirical Application*

In the foregoing sections, three basic elements have been distinguished: soft data, causal relations and graph theory. Graph theory provides a useful framework for causality analysis in the case of soft data. This approach will be illustrated by means of a simplified static causal model for regional (provincial) economic linkages in the Netherlands (1970). This model, partly based on hard information and partly based on soft information, can be represented by means of the following diagram:

The model represented by the arrow scheme in Figure 1 can be described by the following equations:

— a definitive relationship for the activity rate (on a metric scale):

$$a = e/p \tag{18}$$

— an explanatory relationship for socio-geographical attractiveness, defined by regional agglomeration, accessibility, educational facilities and centrality (the four explanatory variables are measured on an ordinal scale):

$$s = f(l_1, l_2, l_3, l_4) \tag{19}$$

— a quasi-production function which relates regional product to infrastructure capital and the activity rate:

$$p_r = \alpha_0 i^{\alpha_1} a^{\alpha_2} \tag{20}$$

— an income formation relationship that relates average regional income to average gross regional product, average regional wealth and the sociographical attractiveness (on a metric scale; see Nijkamp, 1981):

$$y_r = \kappa_0 + \kappa_1 p_r + \kappa_2 w + \kappa_3 s \tag{21}$$

First, attention will be focused on the causality structure of this model. Figure 1 can be regarded as a directed graph $G_d = (z, t)$ with a set of vertices $z = \{a, s, p_r, y_r\} \cup \{p, e, w, i, l_1, l_2, l_3, l_4\}$ and edges (the arrows in Figure 1) $t = \{(z_i, z_j) \mid (z_i, h_i) \wedge (h_i, z_j) \in m\}$. Given the partition $z = y \cup x$, as well as a set of binary relations of the form $h_i R z_j$, $h_i \in H = \{h_1, h_2, h_3, h_4\}$, the *causal structure*

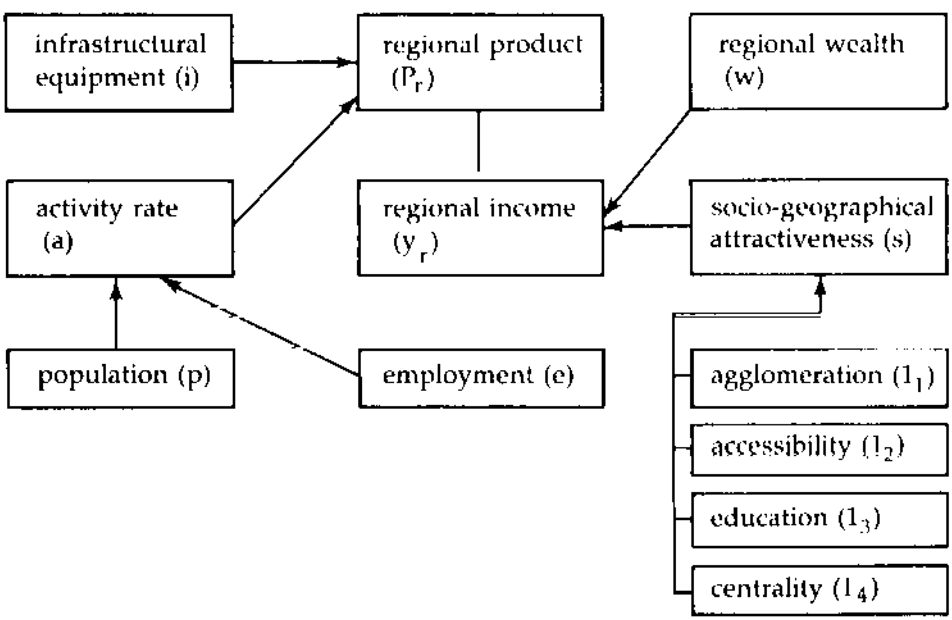


FIGURE 1. Causal Flow Diagram of a Simple Regional Income — Formation Model

of the spatio-economic system (18)-(21) can be expressed in the form of the adjacency matrix A (cf. section 2):

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & | & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & | & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & | & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (22)$$

Replacement of the non-zero elements of A by the 'expected' sign of the coefficients in system (18)-(21) may easily yield the sign matrix S (see section 2).

Further steps in the empirical application are the calculation of the causality degree, a discussion of the kinds of causality in Figure 1, and a calculation of the α-, β-, γ-, and μ-indices (see section 3). This analysis can be carried out on the basis of the information contained in G_d, so that (even in the case of soft data input) causality analysis is a relevant approach.

According to section 3, the causality degree between the successive variables can be calculated via the degree of the associated vertex between each pair of variables. The number of edges incident on a vertex z_i is called the degree (valency) of the vertex, and is denoted by $d(z_i)$. This can easily be incorporated in Table 1, where the entries indicate the causality degree between each pair of variables.

Concerning the kinds of causality (direct, immediate and mutual), the following comments can be made:

Direct causality relationships, $z_i D z_{i'}$, are defined by pairs of arcs $\{z_i, z_{i'}\}$ ϵ of G_d, viz., $\{a, p_r\}$, $\{i, p_r\}$, $\{w, y_r\}$, $\{p_r, y_r\}$, etc.

Immediate causality relationships, $z_i C z_{i'}$, are defined by means of arcs in which all vertices are distinct (i.e., by paths of the graph G_d), namely $\{(p, a), (a, p_r), (p_r, y_r)\}$, $\{(e, a), (a, y_r), (p_r, y_r)\}$, $\{(l_1, s), (s, y_r)\}$, etc.

Mutual causality relationships, $y_r D y_{r'}$, do not occur in system (18)-(21), since the regional development model possesses a *recursive* causal structure.

TABLE 1. Table with Interrelations and Causality Degrees

	a	s	p_r	y_r	p	e	w	i	l_1	l_2	l_3	l_4	$d(z_j)$
a	–	0	1	0	1	1	0	0	0	0	0	0	3
s	0	–	0	1	0	0	0	0	1	1	1	1	5
p_r	1	0	–	1	0	0	0	1	0	0	0	0	3
y_r	0	1	1	–	0	0	1	0	0	0	0	0	3
p	1	0	0	0	–	0	0	0	0	0	0	0	1
e	1	0	0	0	0	–	0	0	0	0	0	0	1
w	0	0	0	1	0	0	–	0	0	0	0	0	1
i	0	0	1	0	0	0	0	–	0	0	0	0	1
l_1	0	1	0	0	0	0	0	0	–	0	0	0	1
l_2	0	1	0	0	0	0	0	0	0	–	0	0	1
l_3	0	1	0	0	0	0	0	0	0	0	–	0	1
l_4	0	1	0	0	0	0	0	0	0	0	0	–	1

The topological measures for graph structure can also be calculated according to the indices discussed in section 3.

The *cyclomatic* number μ (corresponding to the network shown in Figure 1) is equal to zero. This means that the recursive causal system (18)-(21) can be viewed as a branching network. From definition (15), it can be derived that: $\mu = 0 - \alpha = 0$. This indicates a 'minimal spanning tree' in which the removal of any edge $(z_i, z_j) \epsilon\ I$ would break G_d into two unconnected subgraphs.

Calculation of the *beta*-index yields the value $\beta = .92$, indicating a very simple network structure (the lower portion of the β-scale — i.e., from zero to one — differentiates between different types of branching networks). The low value of the *gamma-index*, $\gamma = .37$, confirms the fact that the causal structure of the regional development model is 'minimally' connected, indicating a poorly interwoven spatial system.

The foregoing causality analysis demonstrates that meaningful conclusions can be inferred regarding the causal structure of an interrelated soft economic model by using notions from graph theory.

The second part of the empirical application concerns the estimation of equations (18)-(21), taking into account the non-metric nature of (19). The regional structural model can be expressed in the following stochastic form:

$$Y\Gamma + XB = U \tag{23}$$

where $Y = [y_1, y_2 \ldots y_r \ldots y_R]'$ is a $(R \times N)$ matrix with, as elements, the $(1 \times N)$ vectors with data y_r on all N endogenous variables in all regions r; Γ an (NXN) matrix with coefficients for endogenous variables; X a (RXK) matrix with, as elements, the $(1XK)$ vectors with data x_r on all K exogenous variables in all regions r; B a (KXN) matrix with coefficients of exogenous variables; and U the (RXN) matrix, each row of which is the $(1XN)$ vector $u_r \sim NID\ (0, \Sigma)$ with error terms.

Since Σ is a triangular matrix and the variance-covariance matrix $\Sigma = E(u_r' u_r)$ has a diagonal form, the simultaneous-equation system (23) is a *recursive* equation system. This means that (in principle) each equation of system (23) may be estimated by OLS. Equation (18) is a definition and can be calculated on the basis of available (metric) regional cross-section data. Equation (20) is a (metric) quasi-production function of a Cobb-Douglas type, and can easily be estimated given the availability of data on p, i and a (see Nijkamp, 1982). The results for the coefficients are:[1]

$$\begin{aligned} \ln \hat{\alpha}_0 &= -0.973\ (0.750) \\ \hat{\alpha}_1 &= \ \ 0.620\ (0.240) \qquad R^2 = 0.515 \\ \hat{\alpha}_2 &= \ \ 1.017\ (0.578) \end{aligned}$$

Equation (21) is a linear relationship which can readily be estimated via OLS (see Nijkamp, 1981). The results are:[2]

[1] Figures in parentheses are standard deviations.
[2] Figures in square brackets are t-values.

$$\hat{\kappa}_0 = 2947.30\ [5.46]$$
$$\hat{\kappa}_1 = 0.02\ [2.90]$$
$$\hat{\kappa}_2 = 0.28\ [4.84] \qquad R^2 = 0.928$$
$$\hat{\kappa}_3 = 137.07\ [1.29]$$

The assumption of (21) was that s is a metric variable. In reality, however, the socio-geographical attractiveness s is a *qualitative* variable which cannot be directly measured on a cardinal scale. It is composed of soft data on l_1, l_2, l_3 and l_4. These regional indicators have been measured as regional *ordinal* data. In order to use s in equation (21), it was necessary to transform the ordinal information on l_1, l_2, l_3 and l_4 into cardinal information via a multidimensional scaling analysis (see section 1). It turned out that a one-dimensional configuration led to a fairly high goodness-of-fit for the transformation of the ordinal information into a cardinal configuration. Thus the sequence of operations carried out was: ordinal data on $l_1, \ldots, l_4 \rightarrow$ multidimensional scaling of ordinal data $\rightarrow$ cardinal one-dimensional configuration for regional socio-geographical attractiveness $\rightarrow$ use of provincial configuration of s in (21) (see Nijkamp, 1981).

The latter part of this analysis shows that methods for dealing with nonmetric variables may form a useful complement to traditional techniques.

The conclusion from this last part is that soft econometric techniques may be useful tools for dealing with qualitative information in causal regional modelling. Of course, it should be borne in mind that the foregoing model has been illustrative, and that more extensive research into soft causal models is needed.

REFERENCES

Adelman, I. and Morris, C. T. 1974. The derivation of cardinal scales from ordinal data. In *Economic Development and Planning*, ed. W. Sellebaerts, pp. 1-39. London: Macmillan.

Andrasfai, B. 1977. *Introductory Graph Theory*. Bristol: Hilger.

Basman, R. L. 1963. The causal interpretation of nontriangular systems of economic relations. *Econometrica* 31: 439-448.

Behzad, M. and Chartrand, G. 1971. *Introduction to the Theory of Graphs*. Boston: Allyn and Bacon.

Beineke, L. W. and Wilson, R. J., eds. 1978. *Selected Topics in Graph Theory*. London: Academic Press.

Bennett, R. J. 1979. *Spatial Time Series*. London: Pion.

Bennett, R. J. and Chorley, R. J. 1978. *Environmental Systems: Philosophy, Analysis and Control*. London: Methuen and Co.

Blalock, H. M. 1964. *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

Blommestein, H. J. 1981a. The detection and representation of causal orderings in spatial economic systems. The Netherlands: Twente University of Technology, Department of Public Administration, Working Paper No. 1981-6.

Blommestein, H. J. 1981b. Alternative approaches to spatial autocorrelation: a further improvement over current practice. The Netherlands: Twente University of Technology, Department of Public Administration, Working Paper No. 1981-5.

Blommestein, H. J. and van Deth, J. W. 1981. The analysis of soft information for urban planning processes. Paper presented at the 8th European Symposium on Urban Data Management, Oslo, Norway.

Blommestein, H. J. and Nijkamp, P. 1981. *Soft spatial econometric causality models.* Amsterdam, The Netherlands: Free University, Department of Economics, Research Memorandum 1981-20.

Christofides, N. 1975. *Graphy Theory; An Algorithmic Approach.* New York: Academic Press.

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

Domencich, F. A. and McFadden, D. 1975. *Urban Travel Demand, a Behavioral Analysis.* Amsterdam: North-Holland.

Folmer, H. 1980. Measurement of the effects of regional policy instruments. *Environment & Planning A* 12: 1191-1202.

Fox, K. A., Sengupta, J. K. and Thorbecke, E. 1966. *The Theory of Quantitative Economic Policy.* Amsterdam: North-Holland.

Garrison, W. L. and Marble, D. F. 1962. *The Structure of Transparation Networks.* U.S. Army Transportation Command, Technical Report 62-11.

Gilli, M. 1980. CAUSOR — a program for the analysis of recursive and interdependent causal structures. Université de Genève, Départment d'économetric, User's Manual.

Grizzle, J. E., Starmer, C. F. and Koch, G. G. 1969. Analysis of categorical data by means of linear models. *Biometrics* 25: 489-504.

Harary, F., Norman, R. Z. and Cartwright, D. 1965. *Structural Models: An Introduction to the Theory of Directed Graphs.* New York: John Wiley.

Harvey, D. 1969. *Explanation in Geography.* London: Edward Arnold.

Hepple, L. W. 1976. A maximum likelihood model for econometric estimation with spatial data. In *Theory and Practice in Regional Science,* ed. I. Masser. London: Pion.

Hordijk, L. 1979. Problems in estimating econometric relations in space. *Papers of the Regional Science Association* 42: 99-118.

Jöreskog, K. G. 1977. Structural equation models in the social sciences. In *Applications of Statistics,* ed. P. R. Krishnaiah, pp. 265-286. Amsterdam: North-Holland.

Kansky, K. J. 1963. *Structure of transport networks: relationships between network geometry and regional characteristics.* Chicago: University of Chicago, Department of Geography, Research Paper.

Lazarsfeld, P. F. 1954. Interpretation of statistical relationships as a research operation. In *The Language of Social Research,* eds. P. F. Lazarsfeld and A. Rosenberg, pp. 115-125. Glencoe: Free Press.

Lehnen, R. G. and Koch, G. P. 1974. A general linear approach to the analysis of nonmetric data. *American Journal of Political Science* 18: 283-313.

Leitner, H. and Wohlschlägl, H. 1980. Metrische und ordinale pfadanalyse. *Geographische Zeitschrift* 68: 61-106.

Marshall, C. W. 1971. *Applied Graph Theory.* New York: John Wiley.

McCullagh, S. 1980. Regression models for ordinal data. *Journal of the Royal Statistical Society* 42: 109-142.

Nijkamp, P. 1979. *Multidimensional Spatial Data and Decision Analysis.* Chichester, New York: John Wiley.

Nijkamp, P. 1980. *Environmental Policy Analysis.* Chichester, New York: John Wiley.

Nijkamp, P. 1982. A multidimensional analysis of infrastructure and regional development. In *Planning of Structural Economic Change in Space and Time,* ed. A. Andersson. Amsterdam: North-Holland.

Nijkamp, P. and Rietveld, P. 1980. *Ordinal multivariate analysis.* Laxemburg: IIASA, Professional Paper PP-81-2.

Nijkamp, P. and Rietveld, P. 1982. Causality structures in multiregional economic models. In *Practice of Multiregional Modelling*, eds. P. Nijkamp and P. Rietveld. Amsterdam: North-Holland.

Ord, J. K. 1975. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70: 120-126.

Reitveld, P. 1981. *Causality structures in multiregional economic models.* Laxenburg, Austria: IIASA, Working Paper WP-81-50.

Riguet, J. 1980. A graph-theoretic model for the teaching of some basic concepts in environmental protection. Paper presented at the Education and Environment Seminar, Budapest, Hungary.

Samuelson, P. A. 1947. *Foundations of Economic Analysis.* Harvard: Harvard University Press.

Simon, H. A. 1953. Causal ordering and identifiability. In *Studies in Econometric Method*, eds. W. Hood and T. C. Koopmans, pp. 49-74. New York: John Wiley.

Simon, H. A. 1954. Spurious correlation: a causal interpretation. *Journal of the American Statistical Association* 49: 467-479.

Simon, H. A. 1957. *Models of Man.* New York: John Wiley.

Theil, H. 1971. On the estimation of relationships involving qualitative variables. *American Journal of Sociology* 76: 103-154.

van Lierop, W. G. J. and Nijkamp, P. 1980. Spatial choice and interaction models: criteria and aggregation. *Urban Studies* 17: 299-311.

Wold, H. 1975. Soft modeling by latent variables. In *Perspectives in Probability and Statistics*, ed. J. Gani, pp. 117-142. London: Academic Press.

Wold, H. 1954. Causality and econometrics. *Econometrica* 22: 162-177.

Wrigley, N. 1980. Categorical data, repeated-measurement research designs, and regional industrial surveys. *Regional Studies* 14: 455-471.

Zadeh, L. A., Fu, K. S., Tanaka, K. and Shimura, M. 1975. *Fuzzy Sets and Their Applications to Cognitive and Decision Processes.* New York: Academic Press.