

SPECIFICATION AND ESTIMATION OF SPATIAL ECONOMETRIC MODELS

A Discussion of Alternative Strategies for Spatial Economic Modelling

Hans J. BLOMMESTEIN*

Twente University of Technology, 7500 AE Enschede, The Netherlands

Received January 1982, final version received April 1982

The semantical insufficiency of (spatial) economic theories necessitates the making of additional assumptions — thereby introducing substantial specification uncertainty — in order to arrive at a fully specified econometric model. The traditional or current approach to econometric modelling treats specification uncertainty inadequately. This proposition is illustrated by two well-known examples from the spatial economic literature. Two alternative specification strategies for spatial economic modelling — designed to improve the current spatial econometric modelling approach — are proposed. One of these strategies is used for a specification analysis of agricultural output in Eire.

1. Introduction

Ideally, spatial economic theories should provide the researcher with sufficient a priori information to enable the construction of *fully specified* spatial econometric models. In such a situation the researcher can make an unambiguous choice from a wide range of possible model specifications and appropriate — i.e., in accordance with criteria such as unbiasedness, consistency, efficiency, etc. — econometric/statistical methods.

Unfortunately, this is not the common situation in (spatial) economics [cf. Blommestein (1981a, b), and section 2]. In the natural sciences, physical theories often suggest, for example, detailed forms for correlograms, spectra and frequency responses, which may be compared against empirical data [cf. Heine (1955), Whittle (1954, 1956, 1962)]. Hepple (1974) remarked like we did for (spatial) economics, that in human geography these theories rarely exist.

As a consequence, researchers from the social sciences are confronted with substantial *specification uncertainty*. In order to arrive at a fully specified

*The author is indebted to P. Nijkamp, F.C. Palm and N.V. Ponsen for helpful comments and suggestions on earlier versions of this manuscript. The comments of the referee are also gratefully acknowledged. Any remaining errors are those of the author.

econometric model, the so-called traditional approach to econometric model-building employs to a large extent informal or judgemental information, non-independent pre-tests, etc. [see Zellner (1979), and section 2]. The large-scale, non-systematic use of informal information in the modelbuilding process inspired several authors to harsh methodological criticism of the traditional (or current) econometric modelling approach [see for example Sargent and Sims (1977), Sims (1980)]. In section 3 some examples are given of the traditional approach to econometric modelling in spatial economics [see also Blommestein (1981a, b)].

The proposed specification methodologies — designed to improve the current econometric modelling approach — differ quite substantively from each other. These differences can be partly explained by the treatment of different specification aspects (dynamic structure, direction of causality, functional form, structural/regression equations, etc.) and partly by a different (implicit or explicit) methodological view on econometric model building.¹ This last matter will be further elaborated by a discussion of two of these specification methodologies, while concentrating on the specification of the *spatial dynamics* in a number of typical linear spatial economic models in regression equation form.

For this reason I shall present *spatial versions*² of common factor analysis (COMFAC) and economic factor analysis (ECONFAC), put forward respectively by Sargan (1975, 1980a), and Blommestein and Palm (1980) in a time-domain context. The justification for this are the distinct technical and methodological problems connected to spatial models in general [cf. Cliff and Ord (1975, 1981), Bennett and Chorley (1978), Bennett (1979)], and the modelling of spatial dynamics in particular (see section 4). The latter aspect will be further illustrated by presenting some empirical results in section 5.

2. The nature of spatial economic hypotheses and its consequences for the construction of spatial econometric models

In the introduction it was ascertained that (spatial) economic theories generate insufficient information for a complete specification of econometric relationships. It is of fundamental importance, therefore, to have a proper understanding of the relationship(s) between the nature of (spatial) economic

¹Thus the recent proposed specification strategies differ from each other in two respects: (i) Treatment of different specification aspects, such as (within the context of multi-equations temporal systems) testing the lag length, stability of parameters, and block exogeneity [Sims (1980)], testing overidentifying restrictions [Hendry and Anderson (1977)], checking the implications of restricted structural models for transfer functions and final equations [Zellner and Palm (1974, 1975), Palm (1977)], testing inter — and intra — equation restrictions [Sargent (1981)]. This article focusses on the specification of single-equation *spatial* systems [see Blommestein (1982) for the specification of spatio (-temporal) multi-equations systems]. (ii) A different methodological view on econometric modelbuilding (see section 4 and footnote 10).

²*Spatio-temporal* versions are considered in Blommestein (1981a).

hypotheses and its consequences for the specification of (spatial) econometric models.

Basic economic theories describe so-called generic structures, i.e., a general description of social (economic) structures by a set of concepts, relations and the like, in logical form. Next, with the aid of Hempel's 'bridge principle', the concepts, relations, and the like from basic economic theories are 'translated' for the construction of specific economic theories. The bridge principle of Hempel (1952, 1966) indicates how the fundamental entities and processes (formulated according to Hempel's internal principle) of a basic theory are connected with its corresponding empirical domain. Hempel (1952, 1966) notes that an operationalization of basic theories is necessary to enable the testing of those theories [cf. also Samuelson (1947)].

However, unlike the specific theories in the natural sciences, a specific economic theory is more than just a testable representation of a basic economic theory [Klant (1979)]. As a result of the lack of restrictions (due to the presence of, among others, so-called general *ceteris paribus* clauses) it is possible to deduce so many interpretations that a falsification of basic economic theories becomes very difficult. Papandreou (1958) calls these theories *semantically insufficient*. A specific economic theory is, therefore, in fact an *augmented* theory [Papandreou (1958)], i.e., an interpretation (operationalization) of a basic theory consisting partly of the specification (according to Hempel's bridge principle) of relations, concepts and the like from a basic theory, and partly of additional assumptions with respect to functional forms, lag structures, stochastic properties, the classification of variables into endogenous and exogenous variables, the direction of causality, and the like.

In the words of J.S. Cramer (1969, p. 2): 'Unfortunately economic theorists set great store by generality, and their models are therefore as a rule insufficiently specific to permit an empirical application. As a consequence, virtually all econometric studies add specific hypotheses of their own which are appropriate to the particular situation under review. The convenient approximations are dictated by the requirements of statistical estimation; they are based on common sense rather than on abstract economic theory.'

The semantical insufficiency of (spatial) economic theories necessitates the making of additional assumptions — thereby introducing substantial specification uncertainty — in order to arrive at a fully specified econometric model. The current or traditional approach to econometric modelling 'solves' this kind of specification uncertainty rather informally (compare the aforegiven citation from J.S. Cramer's book), i.e., a priori information, information obtained from non-independent (pre-)tests, etc. are employed in an ad hoc, unsystematic fashion. According to Zellner (1979), traditional econometric analyses like many statistical analyses, tend to concentrate attention mainly on given models — thereby implicitly ignoring a great deal

of specification uncertainty — and relatively little on systematic (=formal) methods for checking whether formulated models are consistent with information in sample data and for improving (i.e., repairing defects of) proposed models.

The methodology of the traditional approach to econometric modelling is conveniently summarized by Palm (1981) as follows [see for more details Zellner (1979) and references therein]:

- The researcher specifies an initial model, M_0 making (rather informally) use of preliminary data analyses, and, a priori information obtained from sources like economic theories, institutional considerations, previous empirical studies, etc.
- Next, an appropriate (i.e., in accordance with statistical criteria such as unbiasedness, consistency and efficiency) estimation procedure is chosen, *provided* M_0 is the true (!) model.
- Estimation results are judged (diagnostic checking phase) on the basis of conventional criteria such as, significance, 'expected' sign and magnitude of parameter estimates; R^2 values; multicollinearity, heteroscedasticity, spatial (-temporal) autocorrelation in the residuals; and forecasting performance.
- M_0 is respecified (variables are dropped and/or added; the choice for an alternative functional form; modification of the error structure) and — subsequently — re-estimated, when estimation results of M_0 are not satisfactory as judged by one or more of the criteria mentioned above. This iterative process continues until a satisfactorily version — M_f say — is obtained.

Zellner (1979) notes that the main problem with the traditional approach to econometric modelling is that in evaluating the adequacy of M_0 — as well as reformulations of M_0 — a good deal of judgement or prior information is employed, usually rather informally. Therefore, the rather informal use of conventional criteria in judging fitted models, may lead to the acceptance of misspecified models.

The absence of systematic (=formal) methods for comprehensively analyzing sample evidence, diagnostic checking results from initial and respecified models, etc., may lead to invalid and/or contradictory results. For example, Palm (1981) points to the fact that *different* final model specifications have been reported in the economic literature, although *similar* data sets and modelbuilding objectives were employed. Further, specification errors vitally and differently affect the large sample properties of alternative estimation and testing procedures [cf. Zellner (1979)]. The procedures chosen in a particular situation from the large variety of mainly asymptotically justified estimation and testing techniques — under the assumption that the

model is correctly specified — may, therefore, possess undesirable statistical properties.

These observations clearly indicate the need for proper formal sequential testing procedures for model construction purposes. Unfortunately, formal procedures which are both comprehensively (i.e., covering the entire modelbuilding process) and operationally, remain to be developed [cf. Zellner (1977, 1979)]. However, some progress has been made by the development of (semi-)formal model specification procedures, which cover different parts of the modelbuilding process. Section 4 provides some details on two of these semi-formal specification procedures in a spatial context. First, however, some examples and consequences of the traditional approach to econometric modelling in spatial economics are briefly discussed in the next section.

3. Examples of the traditional approach to econometric modelling in spatial economics

In this section two well-known examples of modelling spatial phenomena, as reported in the spatial economic literature, will be compared with the description in the previous section of the methodology of the traditional approach to econometric modelling.

3.1. *Transport expansion in developing countries*

Taaffe et al. (1963) argued that the expansion of transportation networks plays an important role in the economic growth of developing countries. They conjectured that expansion is influenced by many specific economic, geographical and social forces. In particular, the relationships between road mileages (Y), population (X_1) and area (X_2) has been investigated for Ghana and Nigeria with the help of regression analysis. The following results were, among others, obtained (standard errors in parentheses):

Ghana

$$\ln Y = 1.27 + 0.52 \ln X_1, \quad R^2 = 0.49, \quad (1)$$

(0.087)

$$\ln Y = 1.73 + 0.25 \ln X_2, \quad R^2 = 0.10, \quad (2)$$

(0.12)

$$\ln Y = 0.16 + 0.63 \ln X_1 + 0.41 \ln X_2, \quad R^2 = 0.76, \quad (3)$$

(0.063) (0.065)

Nigeria

$$\ln Y = 1.02 + 0.58 \ln X_1, \quad R^2 = 0.65, \quad (4)$$

(0.06)

$$\ln Y = 0.43 + 0.76 \ln X_2, \quad R^2 = 0.49, \quad (5)$$

(0.11)

$$\ln Y = 0.01 + 0.45 \ln X_1 + 0.48 \ln X_2, \quad R^2 = 0.82. \quad (6)$$

(0.05) (0.07)

Regressions (3) and (6) were judged, according to 'conventional' criteria as goodness-of-fit and size of standard errors, as the best by both Taaffe et al. (1963), and Cliff and Ord (1981). In addition the residuals \hat{E} from the regressions were tested for first-order spatial autocorrelation using the I statistic and weighted join-count statistics BB and BW [see for details Cliff and Ord (1973, 1981)]. For both Nigeria and Ghana, the high values of R^2 for the multiple regressions and the very small degree of spatial autocorrelation among the residuals, led Cliff and Ord (1981) to conclude 'that these models are acceptable descriptions of the data'.

3.2. Economic effects of road accessibility

O'Sullivan (1969) analyzed some spatial economic effects of road accessibility in Eire. For this purpose an index of arterial road accessibility (ARA) was constructed, defined for a vertex i in a road network of the T class roads in Eire as

$$A(i, r) = \sum_{j=1}^J d_{ij}, \quad (7)$$

where $A(i, r)$ represents the index for County r , J is the number of vertices in the network, and d_{ij} is the distance in miles by road on the shortest path between the i th and j th vertices.

One feature of the Irish economy related by O'Sullivan to the ARA index (X_1) was the percentage of gross agricultural output (in value terms) of each county consumed by itself (Y). O'Sullivan conjectured the hypothesis (model M_0) that remote areas will tend to be more selfsufficient, so that Y and X_1 should be positively correlated. Application of ordinary least squares (OLS) by O'Sullivan supported that hypothesis (see also section 5).

The residuals \hat{E}_0 of the fitted model M_0 , were analyzed by Cliff and Ord (1973, 1981) with the help of the spatial autocorrelation test statistic I [see for details Cliff and Ord (1973, 1981)]. The test indicated significant positive

spatial autocorrelation [see table 8.2 in Cliff and Ord (1981, p. 210)]. In order to break up the pattern of spatial autocorrelation in the residuals Cliff and Ord tried a series of regressions with Y as dependent variable and X_1 (ARA) and X_2 (6 alternative versions were tried), as independent variables (model M_a say). Since the residuals of all these regressions still displayed significant positive spatial autocorrelation, Cliff and Ord (1973, 1981) decided to use a linear regression model (model M_1 say) with an autoregressive model for the error terms [see also Ord (1975)]:

$$Y = X\beta + U_1, \quad \text{and} \quad (8)$$

$$U_1 = \rho W U_1 + E_1, \quad (9)$$

where Y is a $(R \times 1)$ column vector with observations on the dependent variable; X a $(R \times K)$ with observations on K (non)stochastic regressors; β a $(K \times 1)$ vector with to be estimated parameters; U_1 a $(R \times 1)$ vector with error terms generated by a first-order Markov process with to be estimated parameter $|\rho| < 1$; W a $(R \times R)$ weighting matrix³ with elements $w_{rr'} \geq 0$, $\forall r \neq r' \in \{1, \dots, R\}$ and $w_{rr} = 0$, representing the nature of interaction between any pair of spatial units r and r' ; and $E_1 \sim \text{NID}(0, \sigma_{E_1}^2 I_R)$.

By comparing the analyses of both Taaffe et al. (1963) and Cliff and Ord (1981) with Palm's summary of the traditional approach to econometric modelling (see section 2), it can be shown that both analyses fit into that description:

- Specification of an initial model M_0 : Taaffe et al. (1963) choose the eqs. (1) and (4), while Cliff and Ord (1981) started with O'Sullivan's model $Y = \beta_0 \iota + \beta_1 X_1 + E_0$, [$\iota = (1 \dots 1)^T$, $E_0 \sim \text{NID}(0, \sigma_{E_0}^2 I_R)$ and β_0, β_1 parameters].⁴
- Choice of an appropriate estimation procedure, under the assumption that M_0 is the true model: In the first instance, both groups of researchers used OLS.
- R^2 values and/or extent of spatial autocorrelation were used as diagnostic criteria.
- Low R^2 values and/or significant spatial autocorrelation led to respecifications of M_0 : Taaffe et al. (1963), tried several alternative models (including non-log versions).⁵ As has been indicated above, regressions (3)

³Also known as the (first order) spatial lag operator or contiguity matrix [cf. Cliff and Ord (1973, 1981)].

⁴ $\iota = (1 \dots 1)^T$ is a $(R \times 1)$ unity column vector in transposed form; I_R is a $(R \times R)$ identity matrix.

⁵Uncertainty with respect to functional forms can formally be incorporated by using the transformations suggested by Box and Cox (1964). For example, a linear model specified as $Y_r = \sum \beta_i X_{i,r} + U_r$, is transformed into the general form $\lambda_0(Y_r) = \sum \beta_i \lambda_i(X_{i,r}) + U_r$, $r \in \{1, \dots, R\}$, where $\lambda(Z) = (Z^\lambda - 1)/\lambda$, $Z > 0$ is the well-known Box-Cox transformation. Since $\lambda_0, \lambda_1, \dots, \lambda_k$ are to be estimated parameters, the functional form is partly determined by the data. Special cases: (i) $\lambda \rightarrow 1$, i.e., a linear model, (ii) $\lambda \rightarrow 0$, i.e., a log-linear model.

and (6) were accepted as satisfactorily models (i.e., model M_f). Similarly Cliff and Ord (1981) also tried a series of regressions (models M_a and M_1). Although model M_1 [eqs. (8) and (9)] was accepted as final version, there is still some evidence of spatial autocorrelation among the residuals (see section 5 for further details).

Both Taaffe et al. (1963), and Cliff and Ord (1981) used a testing framework — commonly applied in the traditional approach to econometric modelling — what has been called by Mizon (1977) the ‘reverse’ procedure which, starting from the most restricted model (which is often the simplest; in this case model M_0), tests sequentially the need for more general models like M_a and M_1 . This is equal to an extension of the parameter set $\{\gamma_i\}$ in the following way:

$$\begin{array}{ll}
 H_0: & M_0(L_s^{B_0}) \leftrightarrow \gamma_0 = 0, & B_0 = \bar{n}, \\
 H_{01}: & M_1(L_s^{B_1}) \leftrightarrow \gamma_0 \neq 0, \quad \gamma_1 = 0, & B_1 > B_0, \\
 H_{012}: & M_2(L_s^{B_2}) \leftrightarrow \gamma_0 \neq 0, \quad \gamma_1 \neq 0, \quad \gamma_2 = 0, & B_2 > B_1, \\
 \text{-----} & & \\
 H_{012\dots N}: & M_N(L_s^{B_N}) \leftrightarrow \gamma_0 \neq 0, \quad \gamma_1 \neq 0, \quad \gamma_2 \neq 0, \dots, \gamma_{N-1} \neq 0, \quad \gamma_N = 0, & \\
 & & B_N > B_{N-1}
 \end{array} \tag{10}$$

in which hypothesis H_0 indicates that model $M_0(L_s^{B_0})$ is the most simplest (most restricted) model with polynomial matrices $\gamma_j(L_s^{B_{0j}})$ in $L_s^{B_{0j}}$ of orders $B_{0j} \leq \bar{n}$ (\bar{n} is a non-negative lower bound); $L_s^{B_{0j}}$ is a spatial lag operator operating on variable j (see for more technical details section 4); and hypothesis $H_{012\dots N}$ corresponds to the finally retained (i.e., the first hypothesis not rejected) model [in our case the models (3), (6) and M_1]. Naturally, extensions of $M_0(L_s^{B_0})$ [$M_n(L_s^{B_n})$ say] are not necessarily characterized by higher order ($B_n > B_0$) spatial lags, since adding more regressors [like in the models (3), (6) and M_a] also leads to more general models.

It has been noted by Mizon (1977), that a testing framework like (10) has no optimal statistical properties, and has a number of unfavorable characteristics: The absence of an a priori specified maintained hypothesis introduces arbitrary elements (see also section 4) in the testing procedure, which means that is difficult to analyze the statistical power of the reverse procedure. The fact that the test statistics used in a framework like (10) are *not* independent, makes this analysis even more difficult [see Hogg (1961)].

These problems together with other objections to the traditional approach to econometric modelling (see the sections 2 and 4), may lead to acceptance of misspecified models. For this reason alternative strategies for (spatial) economic modelling are considered next.

4. Spatial econometric specification analysis: Some details on specification testing procedures

In section 3 the treatment of specification uncertainty by both Taaffe et al., and Cliff and Ord, were judged as unsatisfactorily. Both approaches tackled specification uncertainty — resulting from the lack of detailed information in, for example, modelling error dynamics and/or systematic dynamics — in accordance with the traditional approach to econometric modelling. Alternatively, more systematic frameworks for specifying spatial economic relationships — viz., spatial versions of common factor analysis and economic factor analysis — will be presented next.

A major objective of both common factor analysis (COMFAC) and economic factor analysis (ECONFAC)⁶ is to mitigate the risk of misspecification due to exclusion of variables and/or long lags when there is insufficient detailed theoretical knowledge about the correct specification, which is — as explained above — the common situation in (spatial) economics. For this reason, both COMFAC and ECONFAC start with a fairly general model (i.e., the maintained hypothesis) such that the ‘true’ model is nested within it. The recommendation to start the specification analysis with a fairly general (=‘unrestricted’) model is motivated further by Zellner and Palm’s (1974) consideration that rejection of a ‘restricted’ model when it is true, will be a less serious error than using a restricted model when the restrictions are imposed incorrectly. They argue that the use of improperly unrestricted models, involves carrying along some extra parameters which may be a less serious problem than the imposition of incorrect restrictions which lead to incorrect values of the parameters. Consider the following class of (linear) dynamic models in regression equation form:

$$Z\gamma(L) = E, \quad (11)$$

in temporal systems $\gamma(L) = (\gamma_0(L), \gamma_1(L), \dots, \gamma_K(L))^T$ denotes a vector of $(K+1)$ scalar polynomials in the temporal lag operator L of orders n_0, n_1, \dots, n_K respectively, with $\gamma_0(L)$ operating on the (normalized) $(T \times 1)$ vector Y with observations on the dependent variable, and $\gamma_j(L)$ on $(T \times 1)$ vectors X_j , $j \in \{1, \dots, K\}$ with regressor variables; $Z = [YX]$; and E a $(T \times 1)$ vector with white noise error terms.

In *spatial* systems formulation (11) may be somewhat more complicated than for the ‘common’ (=temporal) econometric case. This mainly as a result of the increasing complexity specifying higher order (>1) spatial lags in the case of non-binary weights and/or irregular lattices [see Blommestein (1981a,

⁶COMFAC and ECONFAC should not be confused with the well-known factor analysis model as presented in, for example, Harman (1970).

b)]. In temporal systems the n_j th lag is defined as $L^{n_j}X_t = X_{t-n_j}$, whereas in purely spatial systems circular routes $r \rightarrow r' \rightarrow \dots \rightarrow r$ may have to be eliminated. Thus, higher-order spatial lags $L_s^{n_j}$ are obtained by powering the $(R \times R)$ weighting — or contiguity matrix $W(L_s^1 = W)$ and eliminating circular routes. For example, the third order spatial lag, L_s^3 , is defined for the $(R \times 1)$ vector X_j as

$$L_s^3 X_j = C_3 X_j, \quad (12)$$

where $C_3 = (W^3 - \Delta_2 W - W \Delta_2 - \Delta_3 + G_3)S_3$; Δ_2 and Δ_3 are $(R \times R)$ diagonal matrices with elements δ_{2r} and δ_{3r} , corresponding to the leading diagonals of W^2 and W^3 ; G_3 a $(R \times R)$ matrix with typical element w_{rr} , $w_{r'r}$, $w_{rr'}$; and S_3 a $(R \times R)$ diagonal scaling matrix [cf. Cliff et al. (1975) and references therein, for more details].

System (11) can then be reformulated as [Blommestein (1981a)]⁷

$$\gamma(L_s) \text{vec } Z = E_s \quad (11')$$

in which $\gamma(L_s)$ is a $(R \times (R(K+1)))$ super-matrix consisting of $K+1$ $(R \times R)$ polynomial matrices $\gamma_j(L_s^{n_j})$ in $L_s^{n_j}$ of orders n_j , $j \in \{0, 1, \dots, K\}$; $Z = [YX_1 \dots X_K]$ a $(R \times (K+1))$ matrix with spatial observations; $E_s \sim \text{NID}(0, \sigma_{E_s}^2)$ a $(R \times 1)$ vector with error terms; with 'vec' denoting an operation of vectorizing a matrix, stacking column after column.

In both the COMFAC- and ECONFAC framework the values of n_j , $j \in \{0, 1, \dots, K\}$ — denoting the largest lag for each variable — are taken to be sufficiently large that E_s may be treated as white noise. For a priori chosen⁸ maximum polynomial orders \tilde{n}_j for n_j , $j \in \{0, 1, \dots, K\}$ model (11') constitutes the *maintained* hypothesis (i.e., most general hypothesis).

COMFAC considers specializations of the form

$$\rho(L_s^m) \beta(L_s) \text{vec } Z = V, \quad (13)$$

where $\rho(L_s^m)$ is a $(R \times R)$ polynomial matrix in L_s^m of order m ; $\beta(L_s)$ is a $(R \times (R(K+1)))$ matrix consisting of $K+1$ $(R \times R)$ polynomial matrices $\beta_j(L_s^{n_j-m})$ in $L_s^{n_j-m}$ of orders n_j-m , $j \in \{0, 1, \dots, K\}$; and $V \sim \text{NID}(0, \sigma_V^2)$ a

⁷Since the systems (11) and (11') are general initial models with serially and spatially uncorrelated disturbances respectively, the following advantages can — among others — be distinguished [see also Palm's (1981) discussion of temporal systems]: (i) The interpretation of the parameters is facilitated by the inclusion of all the dynamics in the systematic part of the regression equations. (ii) In both (11) and (11') OLS is BLUE when no lagged endogenous variables are present. However, the presence of lagged endogenous variables in (11) means that OLS yields consistent and asymptotically efficient — but biased — estimators, while in (11') OLS gives *inconsistent* estimators.

⁸It is possible, however, to test the lag length in the initial model. For example along the lines as indicated by Sargan (1980b).

$(R \times 1)$ vector with error terms. The factorization in (13) will be valid if $\gamma(L_s)$ in (11') satisfies [cf. Sargan (1975), Blommestein (1981a)]

$$\gamma(L_s) = \rho(L_s^m) \beta(L_s) \quad (14)$$

which implies that the polynomials in $\gamma(L_s)$ have a *spatial common factor* of $\rho(L_s^m)$. It is possible for $\gamma(L_s)$ to have at most p common roots (thus $m \leq p$), viz. $p = \min(n_j)$, $j \in \{0, 1, \dots, K\}$. The presence of m *spatial common factors* in (11') implies K restrictions on the $\gamma_j(L_s^m)$ [compare eq. (14)]. Dividing factorization (13) by $\rho(L_s^m)$ yields

$$\beta(L_s) \text{vec } Z = U \quad (15)$$

which is a linear spatial dynamic model with errors generated by a *spatial* autoregressive process of order m , viz. $U = V/\rho(L_s^m)$. A comparison of (11'), (13) and (15) reveals the 'interaction' between systematic dynamics [$\gamma(L_s)$, $\beta(L_s)$; (= 'explained' part)] and error disturbance properties (= 'unexplained' part) in a dynamic equation [see, e.g., Sargan (1964), Palm (1981)].

The above analysis assumes that the orders of the lag polynomials are known and in practice this will not be the case. Hendry and Mizon (1978, 1980) note, that two formally equivalent two-stage decision procedures might be adopted to solve this problem. However, the empirical implementation of these two approach might lead to different model choices.

One procedure conducts sequential tests for reducing the order of dynamics for all $K+1$ variables simultaneously, i.e., $\hat{n} \leq \tilde{n}$, or for each variable separately, i.e., $\hat{n}_j \leq \tilde{n}_j$, $j \in \{0, 1, \dots, K\}$, until a test statistic (to be discussed below) exceeds the chosen critical value. The second step is to test conditional on the previously determined order of dynamics (\hat{n} or \hat{n}_j) the factorization $\gamma(L_s) = \rho(L_s^m) \beta(L_s)$, using the sequence of ordered and nested hypotheses⁹

$$\gamma(L_s) = \rho(L_s^{\hat{m}}) \beta(L_s), \quad \hat{m} \in \{0, 1, \dots, \hat{p}\}, \quad (16)$$

where $\beta(L_s)$ is a $(R \times (R(K+1)))$ matrix consisting of $K+1$ $(R \times R)$ polynomial matrices $\beta_j(L_s^{\hat{n}_j - \hat{m}})$ in $L_s^{\hat{n}_j - \hat{m}}$ of orders $\hat{n}_j - \hat{m}$, $j \in \{0, 1, \dots, K\}$; $\hat{p} = \min(\hat{n}_j)$.

The alternative procedure first determines how many common factors $\hat{m} \leq \hat{p} = \min(\tilde{n}_j)$, are consistent with the data at the chosen significance level, within the a priori specified maintained hypothesis. Secondly, sequential

⁹Testing for autoregressive polynomials $\rho(L_s^{\hat{m}})$ of orders $\hat{m} = 1, 2, 3 \dots$ implies the testing of one common root at the time and always treating it as real though it might be complex. This may lead to the incorrect conclusion that no common roots are present. This problem can be solved by testing for one *pair* of common factors — i.e., $\hat{m} = 2, 4, 6, \dots$ — at a time, implying in case of complex roots testing for complex conjugate pairs jointly [Sargan (1977)].

testing for zero roots from the set of \hat{m} common roots determined in the first step.

Non-rejection of the common factor hypothesis may lead to restricted models which are non-sensical or difficult to interpret in terms of behavior of economic agents. For example in their study on the demand for money in the Netherlands, Blommestein and Palm (1980) found that non-rejection of the common factor hypothesis lead to a simpler model which was difficult to interpret along steady-state growth paths. For this kind of reasons, Blommestein and Palm (1980) continued their specification search procedure by imposing restrictions which could be interpreted in terms of economic behavior or other a priori considerations. In Blommestein (1981a) I introduced the generic term 'economic factor' analysis for this type of specification search procedures, in order to distinguish them from the statistical-mechanically type of procedures like common factor analysis.¹⁰

ECONFAC considers specializations of the form

$$\gamma^*(L_s) \text{vec } Z = V^* \quad (17)$$

in which $\gamma^*(L_s)$ is a $(R \times (R(K+1)))$ matrix consisting of $K+1$ $(R \times R)$ polynomial matrices $\gamma_j^*(L_s^{n_j^*})$ in $L_s^{n_j^*}$ of orders $n_j^* \leq \tilde{n}_j$, $j \in \{0, 1, \dots, K\}$; and V^* a $(R \times 1)$ vector with error terms, not necessarily white noise.

The restrictions imposed on $\gamma(L_s)$ to arrive at $\gamma^*(L_s)$ are to be based on economic theory considerations such as: non-linear cross-equation restrictions in rational expectations models [cf. Sargent (1981)]; partial adjustment models for endogenous variables [cf. Blommestein and Palm (1980)], 'error correction' mechanisms [cf. Hendry (1978), Blommestein and Palm (1980)] and/or 'integral correction' mechanisms [cf. Hendry and Von Ungern-Sternberg (1979)], exclusion restrictions as the result of some causal mechanism [see Palm (1981), Sargent (1981)], requirements of homogeneity

¹⁰This two-fold classification into procedures which impose statistical-type of restrictions, [for example, Sargan (1975), Hendry and Mizon (1978, 1980), Sims (1980), Sargent and Sims (1977)] and procedures which impose economical — (or other a priori information) type of restrictions [for example Zellner and Palm (1974, 1975), Blommestein and Palm (1980), Sargent (1981)], corresponds, respectively, with measurement without (much) theory and measurement with (much) theory. This classification of procedures is, as virtually any classification, not without problems: (i) Leamer's (1978) approach is, by combining both type of restrictions in a Bayesian framework, of a hybrid nature. (ii) Although Sargan (1975) and Hendry and Mizon (1978, 1980) aim (ex ante) at a parsimonious and *theoretically* plausible parametrization [see eq. (15)] that is *not rejected by the data*, it has been noted in Blommestein and Palm (1980), and Palm (1981) that COMFAC may yield (ex post) common factor restrictions [see eq. (16)] which are theoretically difficult to interpret. (iii) Sims (1980), and Sargent and Sims (1977) favor the generation of restrictions within a Walrasian — hence theoretical — framework. Since these kind of restrictions, usually, leads to very complex models [see, for example, Sargent (1981)] — which are often informationally too demanding — Sargent and Sims prefer to work, instead, which less parsimoniously parameterized (reflecting the absence of *operational* restrictions) multiple-equation systems.

of degree zero or one with respect to some or all explanatory variables yield testable restrictions on the parameters of the initial model M_0 [see Palm (1981)], dynamic long-run properties [see Currie (1981)].

In other words, a priori information derived from economic theories is used to impose economically meaningful restrictions on $\gamma(L_s)$. It can be noted that the maintained hypothesis might be a useful empirical framework for confronting alternative (possibly competitive) economic hypotheses. Although we stressed until now the use of a priori information from economic theories in the (traditional) narrow sense, other sources of a priori information — to obtain meaningful restrictions on $\gamma(L_s)$ to define $\gamma^*(L_s)$ — might (or have to) be used as well. For example, institutional considerations of policy intervention rules, changes of technologies, physical-geographical constraints (e.g., natural barriers), etc., constitute part of the (deterministic/stochastic) environment faced by — and potentially influence therefore the behavior of — private and public agents. Therefore — as has been forcefully stressed by Lucas and Sargent (1980), among others, in discussing specification and estimation of dynamic economic equilibrium models and 'rational expectations' models — instead of merely estimating the parameters of decision rules, the parameters of agent's objective functions and (stochastic) environments should be jointly (thereby imposing testable restrictions on the parameters) estimated.

Specification search procedures like ECONFAC which aim at selecting theoretically meaningful models,¹¹ should — in principle — pay attention to the specification of the error term V^* , within the context of theorizing as well. For example, Hansen and Sargent (1980) developed two different models of the error terms, both of which have a behavioral interpretation. However as has been noted by Sargent (1978), technical requirements of econometric identification and/or estimation, may enforce the researcher to impose restrictions which are partly *not* grounded in economic theory or other a priori information. It can be noted that these situations provide additional specific arguments to conduct econometric modelling within general, maintained hypotheses like (11) and (11').

In discussing ECONFAC, frequent use was made of the term 'interpretation'. The word *interpretation* refers both to the formulation of a correspondence between a definitional system (basic theory) and empirical observations (empirical domain), and the comparison of empirical results (estimation results and the like) with that correspondence (see also section 2). Unfortunately, as a result of the semantically insufficiency of (spatial) economic theories, it is not possible, in general, to establish a one-to-one mapping or correspondence; this in turn may lead to substantial specification uncertainty (see also our discussion in section 2). In this section two procedures were proposed to tackle the problem of specification uncertainty.

¹¹Such as represented by eq. (17).

Both procedures employ a testing framework consisting of sequentially testing hypotheses of the following form:

$$\begin{aligned}
 H_{R_1}: M_{R_1}(L_s^{l_1}) \leftrightarrow \gamma l_1 = 0, & \quad l_1 < \tilde{n}, \\
 H_{R_2}: M_{R_2}(L_s^{l_2}) \leftrightarrow \gamma l_1 = \gamma l_2 = 0, & \quad l_2 < l_1, \\
 H_{R_3}: M_{R_3}(L_s^{l_3}) \leftrightarrow \gamma l_1 = \gamma l_2 = \gamma l_3 = 0, & \quad l_3 < l_2, \\
 \dots & \dots \\
 H_{R_N}: M_{R_N}(L_s^{l_N}) \leftrightarrow \gamma l_1 = \gamma l_2 = \dots = \gamma l_N = 0, & \quad l_N < l_{N-1}.
 \end{aligned} \tag{18}$$

The set with hypotheses (18) represents a sequence of ordered and nested hypotheses, whereby hypothesis H_{R_n} indicates that model $M_{R_n}(L_s^{l_n})$ consists of polynomial matrices $\gamma_j(L_s^{l_{nj}})$ in $L_s^{l_{nj}}$ of orders l_{nj} . This sequence of ordered and nested hypotheses is equivalent to impose restrictions on the set with parameters $\{\gamma_{l_i}\}$ such that, $M_{R_N} \subset M_{R_{N-1}} \dots \subset M_{R_1} \subset M$ (M is the model under the maintained hypothesis). It should be noted that the hypotheses are always tested against the *immediately* preceding hypothesis, and not all against the maintained hypothesis [cf. Mizon (1977)]. Sequential frameworks like (18), i.e., procedures which test sequentially hypotheses in increasing order of restrictiveness, are uniformly most powerful (UMP) in the class of procedures with fix the probabilities of accepting a less restricted hypothesis than the true one [see Anderson (1971)]. Since each stage within the — two formally equivalent — 2-stage COMFAC procedures possesses a structure like (18), the test procedure within each stage is UMP. Although the same arguments are valid for the ECONFAC framework, one may face the additional difficulty of having to choose between 2 or more a priori plausible hypotheses to define $\gamma^*(L_s)$. Seen our discussion in 2 and our observation above on the lack of a one-to-one correspondence, this is not surprisingly of course.

The number of restrictions (π_n say) imposed on $\gamma(L_s)$ to define, for example, factorizations like $\rho_n(L_s^m)\beta_n(L_s)$ or $\gamma_n^*(L_s)$, can be tested by means of the likelihood ratio test (LR), the Wald test (W) or Lagrange multiplier test (LM). When $H_0: \phi_n(\gamma) = 0$ is true — $\phi_n(\gamma)$ is a $(\pi_n \times 1)$ vector with restrictions, written in implicit form — the three test statistics (LR , W and LM) are asymptotically distributed as $X_{\pi_n}^2$ [cf. Sargan (1975) Hendry and Mizon (1978, 1980), among others]. The test statistics can be used for each stage of the COMFAC- or ECONFAC procedure, whereby the set with to be tested hypotheses possesses a structure like (18).

The choice of significance levels is very important in these multistage sequential testing procedures. In order to control the probability of Type I error for a procedure as a *whole*, one can use the Bonferroni inequality to provide a lower bound ε^* on the probability of *not* making a Type I error [cf. Savin (1980)].

5. Spatial econometric specification analysis of consumption of agricultural output in Eire¹²

In section 3 it was shown that Cliff and Ord (1973, 1981) re-analyzed some features of the Irish economy — using O'Sullivan's data — in accordance with the traditional approach to econometric modelling. It has been argued in the sections 2, 3 and 4 that this approach tackles specification uncertainty unsatisfactorily. Alternative (systematic) specification strategies like ECONFAC and COMFAC were proposed. It was ascertained that the choice between COMFAC and ECONFAC is dependent on the (im)possibility of imposing theoretically meaningful restrictions.

Cliff and Ord (1973, 1981) tried to improve O'Sullivan's model (M_0) by the reverse procedure (10). In the first instance they aimed at a 'correctly' specified model (M_a), by extending model M_0 with a measure for the importance of rail transport (in fact a *reverse* ECONFAC analysis). Since the residuals \hat{U}_a of model M_a also exhibited significant positive autocorrelation according to the misspecification test I, Cliff and Ord (1981, p. 214) argued that '... an autoregressive model might be worth considering, since such a model would allow for the persistence of regional variations such as those caused by inertia'. Further they add, that the high level of spatial autocorrelation suggests that '... "home" consumption of agricultural output is determined by historical and *other* (italics mine, HB) factors as well as road accessibility'. [Cliff and Ord (1981, p. 239).] We take these statements as evidence that Cliff and Ord made a shift in their specification search strategy from modifying model M_0 based on (spatial) economic theory arguments (model M_a), to adaptations (model M_1) based on sample evidence (information from raw or transformed data). For this reason we shall re-examine the specifications of some models — based on the O'Sullivan data — reported in Ord (1975), and Cliff and Ord (1973, 1981), within the COMFAC framework.

Since the most common tests of misspecification, such as the I statistic, usually tests for a first-order spatial Markov scheme among the residuals, the detection of significant spatial autocorrelation, usually leads to the formulation of models with first order autoregressive errors. This implies (recall the description of the traditional approach to econometric modelling in section 2) that *untested* common factors are imposed which can lead to acceptance of models with a misspecified dynamic structure. The modification of model M_0 by Cliff and Ord (1973, 1981) [see also Ord (1975)], is an example of this modelling approach, as can be shown as follows. Consider the following spatial interaction model (M_2 say):

$$Y = \gamma_0 I + \gamma_1 X_1 + \gamma_2 W X_1 + \gamma_3 W Y + E_1, \quad (19)$$

where Y , X_1 are $(R \times 1)$ vectors with observations (the O'Sullivan data)

¹²The author is indebted to B. Tummers for computational assistance.

recorded at R ($=26$) counties; W a $(R \times R)$ contiguity matrix ($W = L_s^1$) with elements defined by Cliff and Ord (1973, 1981) as $w_{rr'} = q_r(r')/d_{rr'}$ [$q_r(r')$ is the proportion of the interior boundary of County r which is in contact with County r' , and $d_{rr'}$ is the distance between the (geometric) centers of Counties r and r']; the weights were scaled so that $\sum_{r'} w_{rr'} = 1, \forall r \in \{1, 2, \dots, R\}$; $\iota = (11 \dots 1)^T$ and $E_1 \sim \text{NID}(0, \sigma_{E_1}^2 I_R)$ a $(R \times 1)$ vector with disturbances.

Model M_2 [eq. (19)] is chosen in this study as the maintained hypothesis [compare also eq. (11')]. The choice of an initial model is very important, since even if the specification searches are done efficiently they may result in a poor choice of model if the maintained hypothesis is chosen badly. For this reason the assumptions underlying the estimated initial model should be carefully checked with the help of *misspecification* (diagnostic) tests, like the I statistic, spatial correlogram analysis of the residuals, LM test statistic (the W and LR test statistics are usually associated with specification tests), etc. [see Blommestein and Palm (1980) among others, for an example in testing the adequacy of temporal models]. It is of interest to note that a testing framework like (10) has the structure of a series of misspecification tests, implying an analogy with the sequential use of LM test statistics [see Mizon (1977)].

By imposing the *spatial* common root restriction $\gamma_2 = -\gamma_3\gamma_1$ on model M_2 a specialization like (13) is obtained.¹³ Since this specialization is equivalent to model M_1 of Cliff and Ord (1973, 1981), we are in the position to test their model by calculating [see also Blommestein (1981a)]: (a) the ML estimates for model M_2 and M_1 , (b) the likelihood ratio $\Lambda = L(\hat{\psi})_1/L(\hat{\Omega})_2$, where $L(\hat{\psi})_1$ and $L(\hat{\Omega})_2$ denote the values of the likelihood functions of the models M_1 and M_2 respectively. If $H_0: \gamma_2 = -\gamma_3\gamma_1$ holds, $-2 \ln \Lambda$ is asymptotically distributed as X_π^2 with π ($=1$) degree(s) of freedom. The restriction imposed on model M_2 can also be tested by calculating the Wald test statistic, defined as [see, e.g., Sargan (1975)]: $WT = R\phi(\hat{\gamma})^T V_\phi^{-1}(\hat{\gamma})\phi(\hat{\gamma})$, with $V_\phi(\hat{\gamma}) = (\partial\phi/\partial\gamma)V_\gamma(\gamma)(\partial\phi^T/\partial\gamma)|_{\gamma=\hat{\gamma}}$; $V_\gamma(\gamma)$ the large sample covariance matrix of the parameters of the unrestricted model and $\phi(\hat{\gamma})$ a $(\pi \times 1)$ vector with (non)linear restrictions, written in implicit form, on the regression coefficients of the unrestricted model. If $H_0: \phi(\hat{\gamma}) = \hat{\gamma}_2 + \hat{\gamma}_3\hat{\gamma}_1 = 0$ holds, $WT \approx X_\pi^2$. It can be noted that the use of WT only requires estimation of the unrestricted model, being very appropriate in cases whereby estimation methods more complicated than OLS have to be used (which is especially true for spatial interaction models like M_1 and M_2 !), or multiple optima of likelihood functions can be expected [see Sargan (1977)]. Estimated parameters of the models M_0 , M_1 and M_2 [resulting from a(n) (re)analysis of O'Sullivan's data] are given in table 1.

In section 2 it was ascertained that estimation and testing procedures are chosen under the assumption that the true model is known. In that case it is possible to evaluate the, mainly asymptotic, statistical properties of

¹³The corresponding spatial common factor is equal to $(I_R - \gamma_3 W)$.

Table 1
Estimated spatial interaction models.

Mode	Estimated parameters with standard errors in brackets					Estimation method	Explained variance
	Constant	X_1	WX_1	WY	WU_1		
M_0	-8.49	0.00527 (0.0007)	—	—	—	OLS	0.700
M_1	-1.32	0.0032 (0.00067)	-0.021 (0.0009)	—	0.76 (0.00021)	ML	0.851
M_2	0.63	0.00245 (0.00068)	-0.00265 (0.00213)	1.018 (0.00023)	—	OLS	0.882
M_2	-7.479	0.0026 (0.00065)	0.00036 (0.0006)	0.62 (0.00011)	—	ML	0.891

estimators and testing procedures. For example, it has been shown by several authors [cf. Ord (1975), Hepple (1976)] that for model M_1 [see eqs. (8) and (9)] with ρ unknown, OLS yields inconsistent estimates. Similarly, as has been noted by Whittle (1954), ordinary least squares estimators are inconsistent in case of model M_2 . The performance of estimation and testing procedures in *finite* samples, has to be addressed by results from exact (or approximate) distribution theory or controlled experiments such as Monte Carlo studies. Unfortunately, not many finite sample results are available, especially for the spatial models considered above. It is possible of course, to evaluate estimation (or testing procedures) — given asymptotic results — on the basis of a particular ‘real’ dataset. For example, Ord (1975, p. 124) concludes from a comparison of the performance of OLS and ML for model M_1 , using O’Sullivan’s data, that the marked *differences* between the estimates, indicate the *need* for the use of the ML approach.

Similarly, one might argue that the striking differences (see table 1) in estimated parameters of model M_2 , using both OLS and ML, indicate the need for the use of the ML approach. However, in the first instance both conclusions are — in our opinion — not very meaningful. This proposition is based on the following reasoning: Estimation- and testing procedures are chosen under the implicit assumption that the model under investigation, is the true model. Unfortunately, as has been ascertained above, researchers from the social sciences are confronted with substantial specification uncertainty. In case the model is misspecified, estimation- and testing results (parameter estimates, calculated test statistics, etc.) are distorted by specification errors. In general, the precise (or even approximate) consequences of this distortion are not known for spatial econometric models. Consequently, a comparison of results of different estimation- and testing procedures in case of misspecified models is not very meaningful.

However, application of specification testing procedures like COMFAC and ECONFAC, which aim at a mitigation of the risk of misspecification, may put these comparisons between alternative econometric- and statistical procedures on a more solid basis. Naturally, only in those cases that a certain econometric model is *not* rejected according to *both* specification- and misspecification tests.

Returning to our case-study, it appears that according to both a specification test (*LR*-test) and a misspecification test (*I*-statistic), model M_1 had to be rejected: Calculation of $-2\ln\Lambda$ (see above, this section) yields a significant (at $\alpha=0.01$) value of 12.46. Further, there is some evidence of spatial autocorrelation among the residuals, because $z = [I - \mathcal{E}(I/H_0)] / [\text{var}(I/H_0)]^{\frac{1}{2}} = 2.0$ [see Ord (1975, p. 124)]. Unfortunately, testing for autocorrelation among the residuals of model M_2 also yields an unsatisfactorily result, viz. $z = 2.053$. However, it should be noted that due to the presence of an autoregressive component, the distribution of the test statistic *I* changes when either model M_1 or M_2 is used [cf. Ord (1975), Cliff and Ord (1981)]. Therefore the values of z should be considered as a rough guide only. Nevertheless, model M_2 deserves closer inspection [see Blommestein (1981b) for further details].

6. Summary and conclusions

It was shown that (spatial) economic theories generate insufficient information for the specification of fully specified econometric models. Therefore researchers from the social sciences are confronted with substantial specification uncertainty. It was concluded — illustrated by two well-known examples from the spatial economic literature — that the traditional approach to spatial econometric modelling treats specification uncertainty inadequately.

Therefore, two alternative specification strategies for spatial economic modelling were proposed. One of these strategies was used for a specification analysis of consumption of agricultural output in Eire. It was concluded that the final specification selected by Cliff and Ord (1981) in a reanalysis of O'Sullivan's (1969) data, must be rejected.

Finally it was ascertained that the use of systematic specification search strategies enables a more meaningful comparison of alternative estimation- and testing procedures in case of 'real' datasets.

References

- Anderson, T.W., 1971, *The statistical analysis of time series* (Wiley, New York).
- Bennett, R.J., 1979, *Spatial time series* (Pion, London).
- Bennett, R.J. and R.J. Chorly, 1978, *Environmental systems: Philosophy, analysis & control* (Methuen, London).

- Blommestein, H.J., 1981a, Spatial econometric specification analysis: A discussion of some methodological and technical issues, Working paper no. 1981-3, Department of Public Administration, Twente University of Technology.
- Blommestein, H.J., 1981b, Alternative approaches to spatial autocorrelation: A further improvement over current practice, Working paper no. 1981-5, Department of Public Administration, Twente University of Technology.
- Blommestein, H.J., 1982, Specification and Estimation of multiple-equation spatial systems, Working paper no. 1982-1, Department of Public Administration, Twente University of Technology.
- Blommestein, H.J. and F.C. Palm, 1980, Econometric specification analysis - An application to the aggregate demand for money in the Netherlands, Paper presented at the Econometric Society World Meeting in Aix-en-Provence.
- Box, G.E.P. and D.R. Cox, 1964, An analysis of transformations, *Journal of the Royal Statistical Society, Series B*, 26, 211-243.
- Cliff, A.D. and J.K. Ord, 1973, *Spatial autocorrelation* (Pion, London).
- Cliff, A.D. and J.K. Ord, 1975, Model building and the analysis of spatial pattern in human geography, *Journal of the Royal Statistical Society, series B*, 37, 297-348.
- Cliff, A.D. and J.K. Ord, 1981, *Spatial processes, models and applications* (Pion, London).
- Cliff, A.D., P. Haggett, J.K. Ord, K. Bassett and R. Davis, 1975, *Elements of spatial structure: A quantitative approach* (Cambridge University Press, Cambridge).
- Cramer, J.S., 1969, *Empirical econometrics* (North-Holland, Amsterdam).
- Currie, D., 1981, Some long run features of dynamic time series models, *The Economic Journal* 91, 704-715.
- Davidson, J.E.H. and D.F. Hendry, 1975, *Modelling consumers' expenditure on non-durables in the United Kingdom*, manuscript (London School of Economics, London).
- Hansen, L.P. and T.J. Sargent, 1980, Formulating and estimating dynamic linear rational expectations models, *Journal of Economics Dynamics and Control* 2, 7-46.
- Harman, H.H., 1970, *Modern factor analysis* (University of Chicago Press, Chicago, IL).
- Heine, V., 1955, Models for two-dimensional stationary stochastic processes, *Biometrika* 42, 170-187.
- Hempel, C.G., 1952, *The fundamentals of concept formation in empirical science* (University of Chicago Press, Chicago, IL).
- Hempel, C.G., 1966, *Philosophy of natural science* (Prentice-Hall, Englewood Cliffs, NJ).
- Hendry, D.F., 1978, Predictive failure and econometric modelling in macroeconomics, Discussion paper (London School of Economics, London).
- Hendry, D.F. and G.J. Anderson, 1977, Testing dynamic specification in small simultaneous systems: An application to a model of building society behavior in the United Kingdom, in: M. Intriligator, ed., *Frontiers in quantitative economics*, Vol. 3 (North-Holland, Amsterdam).
- Hendry, D.F. and G.E. Mizon, 1978, Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money function by the Bank of England, *The Economic Journal* 88, 549-563.
- Hendry, D.F. and G.E. Mizon, 1980, An empirical application and Monte Carlo analysis of tests of dynamic specification, *Review of Economic Studies* 47, 21-45.
- Hendry, D.F. and Th. von Ungern-Sternberg, 1979, Liquidity and inflation effects on consumers' expenditure, Discussion paper (London School of Economics, London).
- Hepple, L.W., 1974, The impact of stochastic process theory upon spatial analysis in human geography, in: G. Board, R.J. Chorley, P. Haggett and D.R. Stoddart, eds., *Progress in geography*, Vol. 6 (Edward Arnold, London) 91-142.
- Hepple, L.W., 1976, A maximum likelihood model for econometric estimation with spatial data, in: I. Masser, ed., *Theory and practice in regional science*, *London Papers in Regional Science* 6 (Pion, London) 90-104.
- Hogg, R.V., 1961, On the resolution of statistical hypotheses, *Journal of the American Statistical Association* 56, 978-989.
- Klant, J.J., 1979, *Spelregels voor economen* (H.E. Stenfert Kroese, Leiden).
- Leamer, Edward E., 1978, *Specification searches: Ad hoc inference with nonexperimental data* (Wiley, New York).
- Lucas, R.E. and T.J. Sargent, 1980, eds., *Rational expectations and econometric practice* (University of Minnesota Press, Minneapolis, MN).

- Mizon, G.E., 1977, Inferential procedures in nonlinear models: An application in a UK industrial cross section study of factor substitution and returns to scale, *Econometrica* 45, no. 5, 1221–1242.
- Ord, J.K., 1975, Estimation methods for models of spatial interaction, *Journal of the American Statistical Association* 70, 120–126.
- O'Sullivan, P.M., 1969, Transportation networks and the Irish economy, L.S.E. Geographical Papers no. 4 (Weidenfeld and Nicholson, London).
- Palm, Franz C., 1976, Testing the dynamic specification of an econometric model with an application to Belgian data, *European Economic Review* 8, 269–289.
- Palm, Franz C., 1977, On univariate time series methods and simultaneous equation econometric models, *Journal of Econometrics* 5, 379–388.
- Palm, Franz C., 1981, Structural econometric modelling and time series analysis – Towards an intergrated approach, Research memo., Department of Economics, Free University, Amsterdam.
- Papandreou, A.G., 1958, *Economics as a science* (Lippincott, Chicago, IL).
- Samuelson, P.A., 1947, *Foundations of economic analysis* (Harvard University Press, Harvard).
- Sargan, J.D., 1964, Wages and prices in the United Kingdom: A study in econometric methodology, in: P.E. Hart, G. Mills and J.K. Wittaker, eds., *Econometric analysis for national economic planning* (Butterworth, London).
- Sargan, J.D., 1975, A suggested technique for computing approximations to WALD criteria with application to testing dynamic specifications, Discussion paper (London School of Economics, London).
- Sargan, J.D., 1977, A generalization of the Aitchinson–Silvey–Durbin significance test and its application to dynamic specification, Discussion paper (London School of Economics, London).
- Sargent, T.J., 1978, Estimation of dynamic labor demand schedules under rational expectations, *Journal of Political Economy* 86, no. 6, 1009–1044.
- Sargan, J.D., 1980a, Some tests of dynamic specification for a single equation, *Econometrica* 48, 879–898.
- Sargan, J.D., 1980b, The consumer price equation in the post war British economy: An exercise in equation specification testing, *Review of Economic Studies* 47, 113–135.
- Sargent, T.J., 1981, Interpreting economic time series, *Journal of Political Economy* 89, no. 2, 213–248.
- Sargent, Thomas J. and Christopher A. Sims, 1977, Business cycle modeling without pretending to have too much a priori economic theory, in: Christopher A. Sims, ed., *New methods in business cycle research: Proceedings from a Conference* (Federal Reserve Bank of Minneapolis, Minneapolis, MN).
- Savin, N.E., 1980, The Bonferroni and Scheffe' multiple comparison procedures, *Review of Economic Studies* 47, 255–274.
- Sims, Christopher A., 1980, Macroeconomics and reality, *Econometrica* 48, 1–48.
- Taaffe, E.J., R.L. Morrill and P.R. Gould, 1963, Transport expansion in underdeveloped countries: A comparative analysis, *Geographical Review* 53, 503–509.
- Whittle, P., 1954, On stationary processes in the plane, *Biometrika* 41, 434–449.
- Whittle, P., 1956, On the variance of yield variance with plot size, *Biometrika* 43, 337–343.
- Whittle, P., 1962, Topographic correlation, power-law covariance functions and diffusion, *Biometrika* 49, 305–314.
- Zellner, Arnold, 1977, Comments on time series analysis and causal concepts in business cycle research, in: C.A. Sims, ed., *New methods in business cycle research: Proceedings from a Conference* (Federal Reserve Bank of Minneapolis, Minneapolis, MN).
- Zellner, Arnold, 1979, Statistical analysis of econometric models, *Journal of the American Statistical Association* 74, 628–651.
- Zellner, Arnold and Franz C. Palm, 1974, Time series analysis and simultaneous equation econometric models, *Journal of Econometrics* 2, 17–54.
- Zellner, Arnold and Franz C. Palm, 1975, Time series and structural analysis of monetary models of the U.S. Economy, *Sankhyā, Ser. C*, 37, 12–56.