SERIE RESEARCH MEMORANDA

AN EMPIRICAL GENERAL EQUILIBRIUM MODEL
FOR THE SPATIAL INTERACTIONS OF
SUPPLY, DEMAND AND CHOICE

Aart F. de Vos and Jacob A. Bikker

Research Memorandum 1989-24
June 1989

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
EN ECONOMETRIE
AMSTERDAM

AN EMPIRICAL GENERAL EQUILIBRIUM MODEL FOR THE SPATIAL INTERACTIONS OF SUPPLY, DEMAND AND CHOICE.

By Aart F. de Vos* and Jacob A. Bikker**

We introduce an econometric model which describes flows of goods or persons from multiple origins to multiple destinations with separate but interdependent explanations for total outflow and total inflow per region as well as the allocation. It is derived from a general structural model for demand, choice and supply, involving prices. We focus on the reduced form for quantities, which resembles the "Extended Gravity Model" introduced by Alonso(1978).
We discuss the relevance of the model in several fields. The reduced form has a statistical structure posing new econometric problems; we give a consistent estimation procedure. Our leading example concerns the regional interaction between demand and supply in inpatient hospital care.

*) Department of Econometrics
Free University
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands

**)De Nederlandsche Bank N.V.
Econometric research and Special Studies Department
P.O. Box 98
1000 AB Amsterdam
The Netherlands

The authors would like to thank Jacob J. de Vries for an important comment concerning the estimation method described in Section 5.

1. A system-wide introduction.

This paper introduces an econometric general equilibrium model that is new in several respects. Primarily it describes situations that are common in regional economics, but its scope seems wider. The structure of the model is suited for a variety of general equilibrium situations in cross sections. A feature which is given special attention is the function of -sometimes unobserved- equilibrium prices in the structural model.

The model is best illustrated by the leading example: the regional structure of hospital care. The model explains the number of people admitted in hospitals, discerning the region where they live and the region (or hospital) where they have been admitted. Explanatory variables fall apart into three groups: characteristics of the population (demand), characteristics of the hospitals (supply) and choice factors like travel time and distances. A general equilibrium approach is appropriate here. An equilibrium model normally involves prices. However, as insurance coverage is complete in most countries, prices are not a relevant factor. Still equilibrium is reached by mechanisms that operate in a way similar to prices. Waiting lists and admission policies operate as regulating mechanisms. We solve this problem by introducing a structural model involving unknown "prices". Equilibrium conditions lead to a reduced form model for quantities and prices. The reduced form for quantities is the empirical model, and involves the three types of explanatory variables mentioned before.
The necessity of this general equilibrium approach is strongly suggested by global empirical evidence. In Bikker(1982) and De Vos and Bikker(1984) the flows of patients from municipalities to hospitals are modelled. Looking at the data for the demand side, the size and characteristics of the population seem to determine the number of hospital admissions. However, looking at the supply side, the capacity of hospitals in terms of beds and specialists seems to do so too. It is not amazing that many hospital care studies only consider either supply or demand, or conclude that "supply creates its own demand". This pitfall can be avoided by using a general equilibrium model with a

rather inelastic demand and supply combined with an elastic hospital choice; so by regional redistribution both sides can come close to their targets.

This empirical problem faced in health care can occur wherever flows of goods or persons from multiple origins to multiple destinations depend on distances and characteristics of origin and destination zones. This system of interactions becomes rather complex when the aggregate outflow per origin zone and the aggregate inflow per destination zone are to be explained as well as the allocation. Standard estimation techniques for cross section data are not applicable. Technical problems arise by the occurrence of multiplicative models (which are standard in cross sections) with additive restrictions. Regional economists denote this situation as "generalized spatial interaction". Examples are: the visit of schools, shopping centres or theatres, commuting in labour markets, trade and migration.

The reduced form for quantities derived from our general equilibrium model has a structure known in regional economics as the "Extended Gravity Model" (EGM), introduced by Alonso(1978) and discussed i.a. in Porell and Hua(1981), Anselin(1982) and Tabuchi(1984). They elaborated the model mainly as a descriptive device. Our version is more in the econometric tradition, having a well defined statistical structure and being based on an economic model that imposes restrictions on the parameter space. In Bikker and De Vos(1980) and Bikker(1982) -unaware of Alonso's contribution- this statistical model was developed and called "three component model" (3CM), as it is based on a decomposition of all variables into supply, demand and allocation aspects in a way similar to the decomposition of disturbance terms in the classical variance components model. We will use the term 3CM specifically for the statistical structure to be described in Section 4. We use the term extended gravity model (EGM) for the wider class of descriptive models with a similar structure.

Our presentation -which is meant to be a description of the econometric issues illustrated by a simple example- starts with the structural general equilibrium model. This is helpful in imposing restrictions on the parameters based on economic theory, and serves as a guide-line in the selection of variables. Moreover it offers a good background to discuss related models.

An important issue is that the structural model includes prices but the reduced form does not. Nevertheless the reduced form is almost completely identified, even to such extent that the equilibrium prices may be calculated. Observed prices might differ from these. This requires some further explanation. In health care there are many equilibrium mechanisms: waiting lists, admission policies, incidentally also prices. Though some mechanisms like waiting lists may be measured, the question is justified whether one can have the illusion to capture the equilibrium forces completely. If this is not the case, wrong conclusions are easily drawn. The concept of an abstract "equilibrium price" is more robust, especially when several instruments operate in a changing composition but with rather constant aggregate effect.

In our model equilibrium prices are implicitly determined by the supply and demand relations. This is a consequence of the assumption that the same structural relations hold for all markets. The role of prices is thus fundamentally different from that in classical demand models.

This feature seems of general importance in modeling economic systems. In small econometric models the simultaneity of quantities and prices is a standard assumption. More complex systems, like Theil's(1980) system-wide approach to microeconomics, are almost invariably demand orientated with exogenous prices. This difference may be explained from statistical features. Observation of prices causes no problems in a two-equation supply and demand schedule, it is even necessary identify the model. In larger systems it is difficult to incorporate supply. There the assumption of price exogeneity is seldom investigated (an exception being Bronsard and Salvas-Bronsard(1984)). This contrasts strongly with mathematical economics where systems with equilibrium prices are the standard approach. The study of hospital care indicates that this approach is very useful, but that it may be misleading to equate observed and equilibrium prices. Thus the study of quantities only gets an extra dimension. To fit known prices into the same structure is an interesting exercise we leave for later studies.

This paper starts with a general set up of the model and a discussion of some applications in Section 2. In Section 3 the structural model is worked out fully. Section 4 introduces the 3CM and its role as reduced form. In Section 5 an estimation method is described. Section 6 gives a simple example of a model for hospital care illustrating the dependencies that are included in the 3CM.

2. The structural model and its applications.

In this section the general aspects of the structural model and the reduced form are discussed. Details -specifically on the stochastic structure- are treated later on.
The structural model is formulated as a demand-choice-supply schedule. Suppose we have m supply markets, denoted by an index j, with supply function

$$(2.1) \qquad S_j = SFM_j \; p_j^{\vartheta}$$

where S is supply, p is a price and SFM a Supply Factor Model: a (multiplicative) model of factors determining supply in equilibrium (optimal capacity utilization). SFM includes a constant and a disturbance term.
Next we have n demand markets, with index i, where aggregate demand obeys

$$(2.2) \qquad D_i = DFM_i \; q_i^{\lambda}$$

Here D is demand, q a price index and DFM the Demand Factor Model.
The allocation is explained by a choice function. X_{ij}/D_i , the "market share" of supply market j on demand market i, is modelled as

$$(2.3) \qquad \frac{X_{ij}}{D_i} = \frac{p_j^{\nu} \; CFM_j \; R_{ij}}{\Sigma_k \; p_k^{\nu} \; CFM_k \; R_{ik}}$$

where CFM, (Choice Factor Model) and R are both models with choice factors.

The index q in the demand equation (2.2) is

$$(2.4) \qquad q_i = \Sigma_j \, p_j^{\nu} \, CFM_j \, R_{ij}$$

the denominator of (2.3). An advantage of this specification is that an elegant model arises; in the next Section it will be compared with similar specifications from the literature.

The equilibrium conditions for the model are

$$(2.5) \qquad S_j = \Sigma_i \, X_{ij}$$

This results in nm+2n+2m equations from which S_j, D_i, X_{ij}, p_j and q_i can be solved as functions of CFM, DFM, SFM and R.

The most important part of the reduced form for quantities are the relations for aggregate demand and supply. They are implicit functions, given in Section 4, which can be written as:

$$(2.6) \qquad D,S = f_{(\delta,\lambda)}(DFM, A^*, R)$$

with

$$A_j^* = SFM_j^{1-\delta} CFM_j^{\delta}$$

$$\delta = \vartheta/(\nu - \vartheta)$$

where the subscripts are deleted to indicate that all variables on the left side depend on all submodels at the right side. Equation (2.6) reveals two important points. First that only the combination $\vartheta/(\nu - \vartheta)$ is identified: we cannot discern supply and choice elasticities but only some mixture. Second, SFM and CFM only occur as the mixture A^*, while DFM occurs unchanged. Thus the reduced form is asymmetric and its interpretation is difficult without an underlying structural model. Restrictions like homogeneity may be formulated for SFM and DFM (in hospital care SFM could be proportional to the number of beds and DFM to

population), but without a structural model it is impossible to recognize these restrictions in the reduced form.
The equilibrium prices can be written as

$$(2.7) \qquad p_j = \{ g_{j(\delta,\lambda)}(DFM,A^*,R).CFM_j \}^{1/\nu}$$

This equation reveals the risk of assuming that observed prices are identical to equilibrium prices. Doing so will in general impose strong restrictions across equations (2.6) and (2.7), as the same variables and disturbances occur in both equations. It is unlikely that models are specified so accurately that these restrictions are fulfilled. And even if they are it is rather complicated to use this information in the estimation process. Anyhow it is clear that incorporating prices in SFM or CFM is not justified. We finally note that it is possible to relate observed prices to equilibrium prices by using (2.7). This may help in identification, specifically of ν which occurs in (2.7) and not separately in (2.6).

Summing up the model structure in economic terms: supply involves prices; demand and choice react on these prices. Large price elasticities in the choice relation and/or small price elasticities of supply and aggregate demand can explain situations where supply and demand at the aggregate level seem to be fulfilled both. Prices are unobserved equilibrium prices, defined by the specification of the relations.

To illustrate the issues raised above, it is useful to consider three fields where EGM-type models have been applied: hospital care, the world trade matrix, and migration.

In clinical health care the structure is rather clear. A choice relation is obviously relevant, but it is also known that where the hospital capacity is higher, the admission rates are higher too: demand is not fixed but reacts on capacity. Supply also counts: hospital managers tend to aim at capacity utilization. Supply and demand are both rather inflexible; the reallocation of patients tends to cushion most of the local frictions between supply and demand. The remaining frictions are reflected in variations in admission rates and capacity utilization rates.

In the literature many models are either supply or demand orientated, which is too restrictive. Models for hospital choice are also popular. Lee and Cohen(1985) for instance suppose that aggregate demand is fixed and only consider choice relations. The implication which is likely to be wrong is that the total number of admissions per hospital is simply the result of demand. In general this will conflict with capacity restrictions. A feedback between demand and supply implies an allocation of patients different from a simple choice relation. In our model this feedback also results in a supply effect on admission rates: they depend on the distance towards, and the capacity of, neighbouring hospitals, as is observed in practice. Note that this interaction can even take place if Feldstein's (1977) conclusion that demand is on the long run infinite holds. The general level of health insurance during a particular year will restrict demand to a level roughly in equilibrium with supply, leaving room for local differences. The cross section study of these differences will reveal the effect of changes in supply and demand when the insurance regulations remain unchanged.

Alonso has formulated his model for the study of migration. Poot(1986) gives an extensive application on inter-urban migration in New Zealand. Apart from the fact that migration is typically a partial adjustment to equilibrium (Nijkamp and Poot(1987)), a fundamental issue is whether it is justified -like Poot does- to use characteristics of the immigration region like the wage level, which is obviously a price. There are three possibilities: this price is exogenous and a) there is no other equilibrium mechanism, or b) it is supplemented by another (unobserved) equilibrium mechanism; c) the wage rate is endogenous. In case a the structure is much simpler than in the EGM; it is made up by equations (2.2) and (2.3), called a two component model in Bikker and De Vos(1980). If prices are endogenous, case c, they may not be included in the reduced form. Only in case b an interpretation may be given to the effects of the exogenous prices in the reduced form, but the situation becomes very complicated.

In Bikker(1982; 1987) trade flows between 80 countries have been explained with the 3CM. From a viewpoint of structural modeling the main contribution of these papers is that two popular approaches in trade theory are encompassed.
Primarily the 3CM is introduced as an extension of the model of Linnemann(1966), who made trade flows dependent on characteristics of the importing and exporting country and distances in a traditional gravity model (i.e. one log-linear regression for all bilateral trade flows, aggregate imports or exports follow by summation).
Deardorff(1984, page 504) states that "The empirical success of equations like that (Linnemann) should be added to our list of phenomena to be explained". Bikker obtains even better empirical results. He shows i.a. that the aggregate trade flows of countries with many close neighbours are grossly overestimated by Linnemanns model. This indicates that at the aggregate level feedback relations exist, which are ignored by Linnemann.
Mainstream models for aggregate trade flows however, as surveyed by Goldstein and Kahn(1985), regard prices as exogenous and are essentially demand models. Here the question is again, as in the case of migration, whether aggregate supply (exports) is simply the total of demand from trade partners or must be considered as a supply function depending on prices. As noted by Van der Meer(1988) the well known demand model of Armington(1969) can be written as our choice equation (2.3) with prices and total demand (imports) considered as being exogenous. Extension of this model with equations for total imports depending on a price index as given in (2.2) and (2.4) and a supply equation like (2.1) gives an explanation of Bikker's results in terms of the Armington model. However, much work remains to be done to reach further unification.

3. The stochastic structure of the model.

The structural model is now considered in more detail. The structure of the submodels, containing parameters and disturbance terms is now explicitly incorporated. The disturbance terms are specification errors, representing unavoidable approximations. So the supply and demand functions become

(3.1) $$S_j = c_s \, SF_j(\psi_s) \, p_j^{\vartheta} \, e^{us_j}$$

(3.2) $$D_i = c_d \, DF_i(\psi_d) \, q_i^{\lambda} \, e^{ud_i}$$

with us_j and ud_i i.i.d. and mutually independent normally distributed errors. The notation $SF_j(\psi_s)$ stands for a (multiplicative) combination of supply factors with parameters vector ψ_s. The assumptions regarding the functional form of the disturbances are the standard ones. The independence of us and ud is in applications like hospital care strongly motivated by the completely different character of the equations (in i.a. trade models this may be questioned). Note however that in the reduced form disturbances will be correlated.
The choice equation deserves some more attention. Let X_{ij} be the demand from market i for supply market j and suppose

(3.3) $$\frac{X_{ij}}{D_i} = \frac{p_j^{\nu} \, CF_j(\psi_c) \, RF_{ij}(\epsilon) \, e^{u_{ij}} \, e^{uc_j}}{\Sigma_k \, p_k^{\nu} \, CF_k(\psi_c) \, RF_{ik}(\epsilon) \, e^{u_{ik}} \, e^{uc_k}}$$

Clearly the restricti.n

(3.4) $$\Sigma_j \, X_{ij} = D_i$$

which we suppose to hold for the data, is taken into account in (3.3). RF_{ij} stands for a (multiplicative) combination of factors representing specific preferences (or resistance) between i and j. A negative power of distance is typical in regional models. The disturbance term u_{ij} completes the model for R_{ij}. We assume the u_{ij} to be independent normally distributed. CF contains the characteristics of the supply markets influencing choice and a disturbance term uc_j (i.i.d) reflects the uncertainty in this part of the model.
Equation (3.3) is well known from the multinomial logit model. In fact it is a limiting case of the modified multinomial logit model (Amemiya and Nold(1975), Parks(1980)). The essential aspect of this "modification" is that in choice models specification errors (errors in

the model for individual choice) are added to the usual uncertainty arising from individual behaviour for given choice probabilities, the "sampling errors". Both sources of uncertainty are usually present, but for grouped data the sampling errors become less important: the variance of a choice fraction like in our case X_{ij}/D_i is p(1-p)/n where n is the group size (D_i) and p the relevant choice probability (the right hand side of (3.3)). Specification errors do not depend on n so for large n they dominate and the influence of sampling errors may be neglected. In the case of hospital choice we have ample evidence that this is justified even at low levels of aggregation (municipalities and hospitals). Extension with sampling errors is feasible but complicates matters too much for the present purpose.
Consequently we consider (3.3) as the complete model.

The crucial assumption to get an elegant and symmetric model is that the index q in the demand equation (3.2) is

$$(3.5) \qquad q_i = \Sigma_j \, p_j^{\nu} \, CF_j(\psi_c) \, RF_{ij}(\epsilon) \, e^{uij} \, e^{ucj}$$

the denominator of (3.3). A similar assumption is made in the nested logit model; McFadden(1978) justifies it by derivation from a structural model containing stochastic utilities. In a natural way the nested logit model leads to (3.2) with q_i defined as in (3.5) being the result of individual choice behaviour. The interpretation that may be given to q_i is, following McFadden, the utility to be derived from the best choice. Total demand depends on this utility. It is noteworthy that q_i is not a price, it depends on prices and other choice factors. This stems to reason: if other determinants than prices are important in making choices, they must also be important in determining the utility to be derived from the best choice.
Another justification of (3.5) is found in nested CES utility functions (Keller(1976)). This justification extends to applications like trade, where the allocation is not based on choice behaviour of individuals.
A question is whether the specification errors should be included in the index. We think they should. In the nested logit model this is obvious: the model including errors stands for the true impulses. The impact of

the stochastic reaction on the impulses due to individual disturbances is assumed to be negligible.

When disturbance terms are introduced the equilibrium conditions for the model

(3.6) $$S_j = \Sigma_i X_{ij}$$

have an important new implication. They must hold exactly (they hold for the data), so when all model equations are substituted into (3.6) a reduction in the dimension of the stochastic process occurs. The dimension of the data is nm (X_{ij}, while S_j and D_i are just sums). We introduced nm+n+2m independent disturbances. As can be seen from (2.6) uc and us can not be discerned if only X, S and D are observed. In Section 5 we will show how the nm+n+m remaining disturbances relate to the reduced form for X_{ij}, which is a well defined model of dimension nm.

4. The reduced form: the three component model.

The reduced form for quantities has the statistical structure that we call three component model. In its simplest form the 3CM is a sophisticated symmetrical model defined as follows:

(4.1) $$D_i = \gamma_0 B_i(\gamma_1) \alpha_i^{\gamma} e^{vi}$$

(4.2) $$S_j = \delta_0 A_j(\delta_1) \beta_j^{\delta} e^{wj}$$

(4.3) $$\alpha_i = \Sigma_j S_j \beta_j^{-1} R_{ij}$$

(4.4) $$\beta_j = \Sigma_i D_i \alpha_i^{-1} R_{ij}$$

(4.5) $$X_{ij} = D_i \alpha_i^{-1} S_j \beta_j^{-1} R_{ij}$$

(4.6) $$R_{ij} = RF_{ij}(\epsilon) e^{uij}$$

with v_i , w_j and u_{ij} independent normally distributed disturbances. $B_i(\gamma_1)$, $A_j(\delta_1)$ and $RF_{ij}(\epsilon)$ are (multiplicative) combinations of exogenous variables with parameter vectors γ_1, δ_1 and ϵ; These parameters together with γ and δ may be estimated (Section 5). γ_0 and δ_0 are parameters as well, but only relevant for the level of the process, that may be estimated separately.

This model occurs, without disturbance terms, in the regional economics literature as the EGM (see references in Section 1). There the indices α and β are called "systemic variables"*). Much attention is devoted to their interpretation. In our view their interpretation can become clear only with a structural model. As defined in (4.3) and (4.4) α and β are just transformations of the dependent variables S and D (using the allocation matrix R); they are intermediate steps to describe the nonlinear model as used in (2.6) and (2.7).

The reduced form for quantities X_{ij} that arises from (3.1) to (3.6) can be written as a 3CM by substituting:

(4.7) $\gamma = \lambda$ $\qquad\qquad$ $\delta = \vartheta/(\nu - \vartheta)$

(4.8) $B_i(\gamma_1) = DF_i(\psi_d)$ $\qquad\qquad$ $A_j(\delta_1) = SF_j(\psi_s)^{1-\delta} CF_j(\psi_c)^{\delta}$

(4.9) $v_i = ud_i$ $\qquad\qquad$ $w_j = (1-\delta)us_j + \delta uc_j$

(4.10) $\alpha_i = q_i$ $\qquad\qquad$ $\beta_j = S_j \, p_j^{-\nu} \, CFM_j^{-1}$

An intriguing fact is that the 3CM is symmetrical in variables (all variables with index i occurring in one equation and all with index j in the other), but not in SFM and DFM. Even if choice is only motivated by the price, (4.2) differs from the supply function (3.1) by a power $1-\delta$, leading to different coefficients for the supply variables. Also the interpretation of the disturbance term differs. Note that uc, the

*) Our notation (also used in Bikker(1984,1987)) is different from that in the regional economics literature due to the fact that our model arose independently and uses different connotations. Notably in the EGM the systemic variables are denoted as D and C, while the notation α and β is used there for parameters which we denote as γ and δ.

disturbances in the specification of CFM, occurs raised to the power δ in the disturbance of (4.2). One combined disturbance term is the result, the original ones cannot be distinguished.

It may seem a waste to define an asymmetrical structural model to interpret a symmetrical reduced form. The applications of the EGM discussed in Section 2 were, with the exception of De Vos and Bikker(1984), not based on a structural model. The models were simply constructed by choosing multiplicative functions of relevant explanatory variables for $B_i(\gamma_1)$ and $A_j(\delta_1)$ without any restrictions. In that case the model is completely symmetric. But there are three reasons for the prominent place we give the structural model in this paper:

- linkage with known models like the multinomial logit in hospital care and the Armington demand model in trade.
- empirical evidence that logical restrictions derived from structural models hold and improve the interpretation. In De Vos e.a.(1983) referrals of patients to specialists are analyzed for a number of specialisms. The estimated elasticity at the demand side, the number of inhabitants, corresponding with B_i in (4.1), is almost unity in all cases. At the supply side the number of specialists, A_j in (4.2), has lower estimated elasticities, about $1-\delta$. In fact this may be seen as a confirmation that patients choose their specialists and not the other way around. In our empirical example we will obtain a similar result.
- application of the 3CM for hospital care for planning purposes may be misleading without the use of a structural model. De Vos and Bikker(1984) try to define the optimal size of hospitals using the model. This leads to unacceptable solutions unless homogeneity restrictions are imposed, which requires a structural model.

5. Statistical aspects of the 3CM.

The 3CM as defined in (4.1) to (4.6) is a rather complex structure. This is mainly due to the interdependent indices α_i and β_j. Given R, they may be computed according to (4.3) an (4.4). It is not allowed however to use them directly for OLS estimation of (4.1) and (4.2) (like

Tabuchi(1984) suggested) as they depend on the disturbance terms. Moreover, the functional dependency has no analytical solution; the model is implicit, nonlinear and singular. In De Vos and De Vries(1988) these problems are discussed in detail, deriving the likelihood function and the FIML estimators; here we present an iterative two stage procedure.

First we have to deal with the dimensions of the model. Together nm+n+m disturbances occur (u_{ij}, v_i and w_j) while we have only nm observations. Reduction to dimension nm is obtained by a decomposition of u_{ij} known from the analysis of variance. We define a function $\tilde{N}$ that transforms the variates u_{ij} to deviations from their means in both directions:

(5.1) $$\tilde{N}(u_{ij}) = u_{ij} - u_{i.} - u_{.j} + u_{..}$$

where a subscript point denotes the average over the corresponding index. Now an equivalent model can be defined which depends on disturbances with the right dimension: $\tilde{N}(u_{ij})$, $v_i-\gamma u_{i.}$ and $w_j-\delta u_{.j}$. To obtain this result we rewrite (4.3) and (4.4) in terms of new indices $\alpha_i\exp(u_{i.})$ and $\beta_j\exp(u_{.j})$, and replace u_{ij} by $\tilde{N}(u_{ij})$ in (4.6), the definition of R. This change in the indices is compensated in (4.1) and (4.2) by a redefinition of the disturbance terms to $\exp(v_i-\gamma u_{i.})$ and $\exp(w_j-\delta u_{.j})$ respectively. Further $\exp(u_{..})$ and the means of $v_i-\gamma u_{i.}$ and $w_j-\delta u_{.j}$ are, together with the constants, included in the level term. That the indices and the disturbances v_i and w_j are redefined influences only the interpretation of the residuals, not the estimation procedure.

The estimation procedure can be decomposed in a similar way. $\tilde{N}(u_{ij})$ can be estimated from the data on the individual flows X_{ij}. Suppose that $RF_{ij}(\epsilon)$ is simply a power of distance (DIS):

(5.2) $$\ln R_{ij} = \epsilon_0 + \epsilon_1 \ln DIS_{ij} + \tilde{N}(u_{ij})$$

then ϵ_1 can be estimated by OLS from

$$(5.3) \qquad \tilde{N}(lnX_{ij}) - \tilde{N}(lnR_{ij}) = \epsilon_1 \tilde{N}(lnDIS_{ij}) + \tilde{N}(u_{ij})$$

as is well known. The resulting estimates of R_{ij} may be used in (4.3) and (4.4) to calculate the indices. The estimate of ϵ_1 is not fully efficient as ϵ_1 enters via the indices into the equations (4.1) and (4.2), but the information on ϵ_1 in the allocation (5.3) is in normal cases dominant by far.

We now concentrate on the more interesting equations (4.1) and (4.2), conditional upon R. Estimation of the coefficients in these equations is not straightforward as the indices α_i and β_j are stochastic and depend on all disturbance terms, according to (4.3) and (4.4). A consistent estimation method is possible approximating the indices in (4.1) and (4.2) by their deterministic parts (as in the two stage least squares procedure) and using these as instrumental variables in the regression. The deterministic parts can only be calculated conditional upon the parameters, so we have to iterate between parameter estimation and calculating instruments (see Hausman, 1983, p439-440).

The estimation procedure becomes:

i- use (5.3) to get estimates of R

ii- do regression on (4.1) and (4.2) with γ and δ restricted to 0 (or other starting values) to get estimates of γ_0, δ_0, γ_1 and δ_1

iii- use the equations

$$(5.4) \qquad \alpha_i = \Sigma_j A_j(\delta_1) \beta_j^{-1+\delta} R_{ij}$$

$$(5.5) \qquad \beta_j = \Sigma_i B_i(\gamma_1) \alpha_i^{-1+\gamma} R_{ij}$$

to get by iteration estimates of the deterministic parts of the indices which serve as instrumental variables

iv- do regression on (4.1) and (4.2) using instrumental variables for the indices to get estimates of γ, δ, γ_0, δ_0, γ_1 and δ_1

v- go back to iii with the new parameter estimates until reasonable convergence is obtained.

This algorithm performs well in practice. It is an improvement on similar algorithms suggested in Bikker and De Vos(1980) and Anselin (1982), that just use the deterministic parts of the indices to replace the true (stochastic) indices. The improvement lies in the use of the deterministic indices only as instruments. The instrumental variable method is required to avoid the "forbidden regression" problem arising in nonlinear models when the deterministic part is simply substituted (Hausman, 1983, note 60)*). Our estimation method is a NL2S (nonlinear two stage least squares) estimator as defined by Amemiya (1983, p366).. The only difference with Amemiya's BNL2S (best NL2S) estimator is that we do not use as instruments the expectations of α and β but their values at zero disturbances.
A slight improvement of the above algorithm is possible adapting step iii. Instead of calculating both instrument vectors in one iteration, they can be calculated separately, including (either in (5.4) or (5.5)) the residuals of the disturbance term not occurring in the equation the instrumental variable is used for. This improves the approximation while the relevant independences are maintained. In the empirical application we will use this method.

6. Micro and macro elasticities; a simple application of the model.

In this section we derive some important features of the 3CM and illustrate these with a simple example. Our application consists of hospital admissions in and from the 11 provinces in the Netherlands *). The model only uses the number of inhabitants (INH), the number of beds (BED) and the distances. We first use the unrestricted 3CM:

---

*) In some examples we recalculated our forbidden regressions using the consistent method. The difference in estimates appeared to be small compared to the standard errors.

*) The data concern compulsory insured only (70% of the population). They come from the yearly survey of the information system of the Dutch national insurance organization, LISZ 1978.

(6.1) $D_i = \gamma_0 \, INH_i^{\gamma 1} \, \alpha_i^{\gamma} \, e^{vi}$

(6.2) $S_j = \delta_0 \, BED_j^{\delta 1} \, \beta_j^{\delta} \, e^{wj}$

(6.3) $\alpha_i = \Sigma_j \, S_j \, \beta_j^{-1} \, R_{ij}$

(6.4) $\beta_j = \Sigma_i \, D_i \, \alpha_i^{-1} \, R_{ij}$

together with (5.2). The estimation results, obtained by the algorithm from the last section, are (standard errors in parentheses):

$\gamma_1 = .92$ (.09) $\gamma = .16$ (.13)

$\delta_1 = .88$ (.07) $\delta = .14$ (.09)

while (5.2) yields an estimate of $\epsilon_1 = -1.96$.
These estimates almost exactly satisfy the theoretical restrictions following from the most plausible structural model. Demand proportionate to inhabitants, supply proportionate to beds and no variables in the choice equation would imply $\gamma_1 = 1$ and $\delta_1 = 1-\delta$. (That the coefficient for distance is close to the value 2 in Newtons gravity model we rather leave to explain to others).
Imposing the restrictions $\gamma_1 = 1$ and $\delta_1 = 1-\delta$ we obtain the following estimates:

$\gamma_1 = 1$ $\gamma = .06$ (.04)

$\delta_1 = .86$ (.08) $\delta = .14$ (.08)

The effect of changes in the explanatory variables can -due to the nonlinear structure of the model- only be calculated exactly by numerical evaluation. Two essential aspects of those effects may be derived analytically: the effect of an everywhere equal percentage change in a variable on all flows, and the matrix representing the differential effect of (small) local changes. The first we call macro elasticities, the second micro elasticities.

The macro elasticities are easily calculated. Suppose a change in the number of beds in every hospital with a factor a , so: $BED_j^* = a.BED_j$; a star denoting the new situation. Then substitution of $\alpha_i^* = a^r.\alpha_i$ and $\beta_j^* = a^s.\beta_j$ into (6.1) to (6.4) makes clear that all equations are fulfilled if $r = \gamma/(\gamma+\delta-\gamma\delta)$ and $s = (\delta-1)/(\gamma+\delta-\gamma\delta)$. This shows that the macro elasticities are constant:

$$(6.5) \qquad \frac{d\ln S}{d\ln BED} = \frac{d\ln D}{d\ln BED} = \frac{\delta_1 \gamma}{\gamma+\delta-\gamma\delta}$$

We call this the macro elasticity of admissions for beds. Our estimate of this elasticity is 0.28. Similarly the macro elasticity for inhabitants is

$$(6.6) \qquad \frac{d\ln S}{d\ln INH} = \frac{d\ln D}{d\ln INH} = \frac{\gamma_1 \delta}{\gamma+\delta-\gamma\delta}$$

Our estimate is 0.72.

The macro elasticities are conceptually the most important outcomes of the model: they have (different from the individual parameters) a clear meaning which is of great interest especially for general policy problems concerning the planning of hospital capacity. The elasticities $\gamma/(\gamma+\delta-\gamma\delta)$ and $\delta/(\gamma+\delta-\gamma\delta)$ function as multipliers of initial supply and demand effects. If the restrictions $\gamma_1 = 1$ and $\delta_1 = 1-\delta$ are imposed according to the structural model, the macro elasticities sum to unity and at the aggregate level a simple model results where admissions are a weighted geometric average of supply and demand factors. It should be noted that the macro elasticities depend crucially on the ratio between γ and δ, which may both be small.

A broader view on the spatial interaction aspects of the model is obtained by considering the effects of small local changes in supply and demand factors. From the model (6.1) to (6.4) differential equations in $d\ln D$, $d\ln S$, $d\ln \alpha$, $d\ln \beta$, dv and dw can be obtained that lead to

$$(6.7) \qquad \begin{bmatrix} d\ln D \\ d\ln S \end{bmatrix} = \begin{bmatrix} I_n-(1-\delta)\gamma Q^{-1}HK' & \gamma Q^{-1}H \\ \delta L^{-1}K' & I_m-(1-\gamma)\delta L^{-1}K'H \end{bmatrix} \begin{bmatrix} dv \\ dw \end{bmatrix}$$

where

$$(6.8) \qquad H_{ij} = S_j \beta_j^{-1} \alpha_i^{-1} R_{ij}$$

$$(6.9) \qquad K_{ij} = D_i \beta_j^{-1} \alpha_i^{-1} R_{ij}$$

$$(6.10) \qquad Q = I_n - (1-\delta)(1-\gamma) HK'$$

$$(6.11) \qquad L = I_m - (1-\delta)(1-\gamma) K'H$$

with I_n and I_m unit matrices of order n and m respectively. This kind of formulae arise from an equilibrium between two linear structures that describe probabilities of going from the demand side to the supply side and vice versa as known from Markov chains and Input-Output models. Table 1 gives the outcomes for the micro elasticities in the present example, offering much information. Enjoying it requires some understanding of the regional structure, which is given in Figure 1. In general the ordering of the regions is such that adherent numbers represent adherent regions, comparable with time series where a strict order of associations exists. Effects depend on adherence and size of the provinces.

The 6th province Utrecht (Ut), that is small and has many neighbours, may serve as an example. The upper left matrix of Table 1 shows that a local increase in the number of inhabitants ceteris paribus causes a large increase in admissions from this region, the elasticity being about 0.9. These additional patients go partly to neighbouring regions,

Table 1 Effects of percentage changes in number of inhabitants (INH) and nummer of beds (BED) on admission from (D) and in (S) 11 provinces. Elasticities x 100.

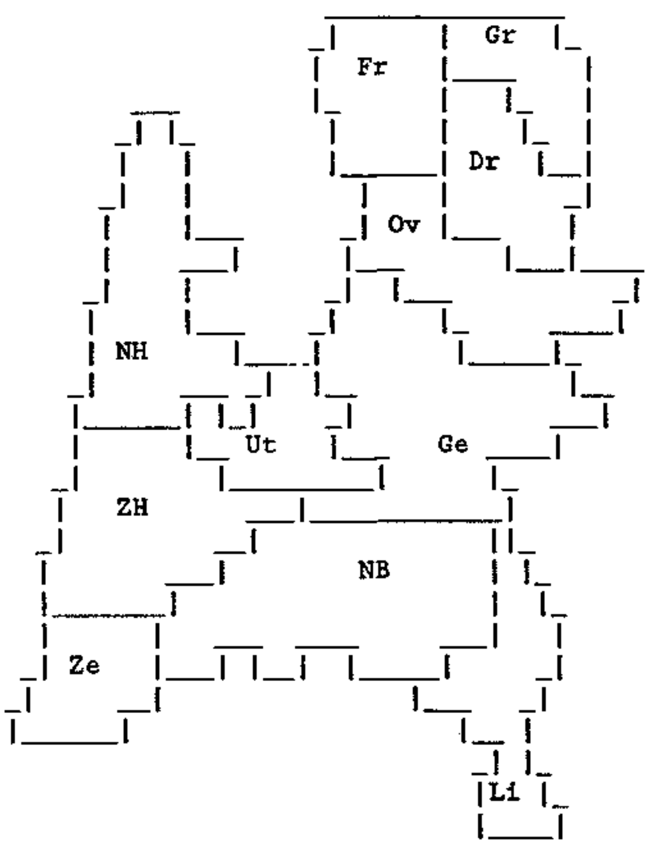
INH

		gr	fr	dr	ov	ge	ut	nh	zh	ze	nb	li
	gr	90	-2	-2	-2	-2	-1	-2	-3	-1	-2	-1
	fr	-2	90	-1	-3	-2	-1	-3	-2	-0	-2	-1
	dr	-2	-2	92	-3	-3	-1	-2	-3	-0	-2	-1
	ov	-1	-2	-1	87	-3	-1	-2	-2	-0	-2	-1
	ge	-1	-1	-1	-2	89	-1	-2	-4	-1	-3	-2
D	ut	-1	-1	-0	-1	-3	91	-4	-4	-1	-3	-1
	nh	-0	-1	-0	-1	-2	-1	85	-4	-1	-2	-1
	zh	-0	-1	-0	-1	-2	-1	-3	85	-1	-3	-1
	ze	-1	-1	-1	-1	-3	-2	-3	-5	94	-4	-2
	nb	-0	-1	-0	-1	-2	-1	-2	-4	-1	86	-2
	li	-1	-1	-0	-1	-3	-1	-2	-3	-1	-4	88
		gr	fr	dr	ov	ge	ut	nh	zh	ze	nb	li
	gr	28	6	4	6	5	2	5	6	1	6	3
	fr	4	30	3	6	6	2	7	5	1	5	3
	dr	4	5	27	7	6	2	5	6	1	5	3
	ov	2	4	3	36	7	3	4	5	1	4	3
	ge	1	2	1	4	33	3	6	8	1	7	5
S	ut	1	2	1	4	7	27	9	10	2	7	3
	nh	1	2	1	2	4	3	41	8	1	5	2
	zh	1	1	1	2	5	3	7	41	2	7	3
	ze	2	2	1	3	6	3	7	12	22	10	5
	nb	1	1	1	2	6	2	5	8	2	39	4
	li	1	1	1	3	6	2	4	7	1	8	36

BED

		gr	fr	dr	ov	ge	ut	nh	zh	ze	nb	li
	gr	13	2	1	2	2	1	2	2	0	2	1
	fr	2	11	1	3	2	1	3	2	0	2	1
	dr	2	2	10	3	2	1	2	2	0	2	1
	ov	1	1	1	14	2	1	2	2	0	2	1
	ge	1	1	0	2	12	2	2	3	0	3	2
D	ut	1	1	0	1	2	11	3	4	0	2	1
	nh	1	1	0	1	2	1	17	3	0	2	1
	zh	1	0	0	1	2	1	3	16	0	2	1
	ze	1	1	0	1	2	2	3	5	7	4	2
	nb	1	0	0	1	2	1	2	3	1	15	2
	li	1	1	0	1	2	1	2	3	0	3	14
		gr	fr	dr	ov	ge	ut	nh	zh	ze	nb	li
	gr	64	-4	-3	-5	-4	-2	-5	-5	-1	-5	-3
	fr	-4	66	-2	-6	-5	-2	-6	-5	-1	-4	-2
	dr	-5	-4	69	-7	-5	-2	-5	-5	-1	-4	-3
	ov	-3	-3	-2	60	-6	-3	-4	-5	-1	-4	-3
	ge	-1	-2	-1	-4	63	-3	-5	-8	-1	-6	-4
S	ut	-1	-1	-1	-3	-6	66	-8	-9	-1	-6	-3
	nh	-1	-1	-1	-2	-4	-3	55	-7	-1	-5	-2
	zh	-1	-1	-1	-2	-4	-3	-6	56	-1	-6	-3
	ze	-2	-1	-1	-3	-5	-3	-7	-11	73	-9	-4
	nb	-1	-1	-1	-2	-5	-2	-5	-7	-1	57	-4
	li	-1	-1	-1	-3	-5	-2	-4	-7	-1	-7	60

Figure 1 The geografical structure of the provinces



where they crowd out a small percentage of the local inhabitants. The particular position of Utrecht is reflected in the fact that additional patients from Utrecht will almost all be admitted: apart from the own hospital capacity there are many nearby provinces to go to.
The matrix down left shows the spread of a possible increase in the number of admissions among the regions of destination. This spread is surprisingly high. The low values of γ and δ are responsible for that: at the aggregate level supply and demand are rather inflexible; changes mainly lead to redistribution of patients. Almost exactly the same situation is encountered in the upper right matrix representing micro elasticities for changes in the number of beds on admissions by origin. Here all elasticities are low as γ is smaller than δ: the supply side is relatively weak in achieving its goals. The matrix right below shows how strong the competition between hospitals is: more beds in one region can lead to a considerable reduction of capacity utilization in neighbouring regions.
The macro elasticities are also incorporated in Table 1: the rows in the left side of the matrix all add to (6.6), in the right side to (6.5). The total matrix is thus singular as it should be because of the restriction that the grand totals of inflows and outflows are equal. The singularity of (6.7) is important for statistical purposes as (6.7) is also the transformation matrix of disturbances to observables (De Vos and De Vries(1988)).

Discerning between macro and micro elasticities is useful for the evaluation of studies that only consider demand or supply. It can be shown that in many of these studies the micro elasticities are estimated directly. The generalization to macro elasticities usually goes wrong. For example: direct estimation of the effect of the number of beds on the number of admissions from regional cross-section data will result in an elasticity about 0.63 (the mean of the diagonal of the matrix down right). The macro elasticity however should be 0.28 according to (6.5). Explaining the number of admissions from both supply and demand factors can lead easily to wrong conclusions because local effects differ strongly from macro effects, due to regional redistribution. The table learns that this can even be the case in this example with a very high level of regional aggregation.

7. Final remarks.

In this paper a newly developed structural model for spatial interaction between supply and demand has been presented. This model gives an economic interpretation to its reduced form, the Extended Gravity Model. A number of econometric issues concerning the statistical version of the EGM, which we call the Three Component Model, are treated. The possibilities of the model are exposed by an application to inpatient health care. Micro and macro elasticities appear to be useful tools to examine the influence of supply, demand and regional redistribution, as well as the influence of the geographical structure on hospital admissions.

The 3CM has been applied to hospital admissions in several other studies (Bikker(1982), De Vos and Bikker(1984)). These applications are based on data about flows of patients from about 200 municipalities to 26 hospitals in the North of the Netherlands (with some variations). Forecasts of the consequences of changes in supply or demand factors appear to be satisfactory. In De Vos and Bikker(1984) it is shown that the effects of closing a hospital can be forecast adequately. In view of these experiences we trust that the model may be useful in other fields as well.

References

Alonso, W.(1978), "A theory of movements". In: Hansen, N.M. (ed.), *Human settlement systems* (Ballinger, Cambridge MA).

Amemiya, T.(1983), "Non linear regression models". In: Griliches, Z. and M.D. Intriligator (eds.), *Handbook of Econometrics vol. 1* (North Holland, Amsterdam).

Amemiya, T. and F.C. Nold(1975), "A Modified Logit Model". *Review of Economics and Statistics*, pp. 255-57.

Anselin, L.(1982), "Implicit Functional Relationships Between Systemic Effects in a General Model of Movement". *Regional Science and Urban Economics* 12, pp. 365-80.

Armington, P.S.(1969a), "A Theory of Demand for Products Distinguished by Place of Production", *IMF Staff Papers*, XVII, pp. 159-78.

Armington, P.S.(1969b), "The Geographical Pattern of Trade and the Effects of Price Changes", *IMF Staff Papers*, XVI, pp. 179-99.

Bikker, J.A. and A.F. de Vos(1980), *Interdependent multiplicative models for allocation and aggregates*. Research memorandum 52, department of econometrics, Free University, Amsterdam.

Bikker, J.A.(1982), *Vraag-aanbodmodellen voor stelsels van geografisch gespreide markten* (demand-supply models for systems of geographically dispersed markets), Free University Press, Amsterdam.

Bikker, J.A.(1987), "An International Trade Flow Model with Substitution: an Extension of the Gravity Model", *Kyklos* 40, pp. 315-37.

Bronsard, C. and L. Salvas-Bronsard(1984), "On Price Exogeneity in Complete Demand Systems". *Journal of Econometrics* 24, pp. 235-47.

Deardorff, A.V.(1984), "Testing trade theories and predicting trade flows". in: Jones, R.W. and P.B. Kenen (eds.), *Handbook of international economics*, Vol. 1 (North Holland, Amsterdam).

De Vos, A.F., T. Melis and J.A. Bikker(1983), *Specialisms and medical consumption: an econometric study of regional supply and demand relations in the Netherlands*, Research memorandum 118, department of econometrics, Free University, Amsterdam.

De Vos, A.F. and J.A. Bikker(1984), *Een structuurmodel voor de klinische gezondheidszorg (a structural model for clinical health care)*, Free University Press, Amsterdam.

De Vos, A.F. and J.J. de Vries(1988), *Full information Maximum Likelihood Estimation of the nonlinear, singular three component model*. research paper, earlier draft presented ESEM 87.

Feldstein, M.S.(1977), "Quality change and the demand for hospital care", *Econometrica* 45, pp. 1681-1702.

Goldstein, M. and M.S. Kahn(1985), "Income and price effects in foreign trade", in: Jones, R.W. and P.B. Kenen (eds.), *Handbook of international economics*, Vol. 2 (North Holland, Amsterdam).

Hausman, J.A.(1983), "Specification and estimation of simultaneous equation models", In: Griliches, Z. and M.D. Intriligator (eds.), *Handbook of Econometrics vol. 1* (North Holland, Amsterdam).

Keller, W.J.(1976), "A Nested CES-type Utility Function and its Demand and Price Index Functions", *European Economic Review* 7, pp.175-186.

Lee, H.L. and M.A. Cohen(1985), "A multinomial Logit Model for the Spatial Distribution of Hospital Utilization". *Journal of Business & Economic Statistics* Vol. 3 n. 2, pp. 159-168.

Linnemann, H.(1966), *An econometric analysis of international trade flows*, North-Holland, Amsterdam.

McFadden, D.(1978), "Conditional Logit Analysis of Qualitative Choice Behaviour". in P. Zarembka (ed.), *Frontiers in econometrics*, (Academic Press), pp. 105-42.

Nijkamp, P. and J. Poot(1987), "Dynamics of generalised spatial interaction models". *Regional Science and Urban Economics* 17, pp. 367-390.

Parks, R.W.(1980), "On the Estimation of Multinomial Logit Models from Relative Frequency Data", *J. Econometrics*, pp. 293-301.

Poot, J.(1986), "A system approach to modeling the intra-urban exchange of workers in New Zealand". *Scottish Journal of Political Economy* 33, pp. 249-274.

Porell, F.W. and C. Hua(1981), "An econometric procedure for estimation of a generalized systemic gravity model under incomplete information about the system". *Regional Science and Urban Economics* 11, pp. 585-606.

Tabuchi, T.(1984), "The systemic Variables and Elasticities in Alonso's General Theory of Movement". *Regional Science and Urban Economics* 14, pp. 249-264.

Theil, H.(1980), *The System-Wide Approach to Microeconomics*. (Basil Blackwell, Oxford).

Van der Meer, T.(1988), *Optimization in econometrics, with economic and statistical applications*, Free University Press, Amsterdam.