A utility framework for interaction models for spatial processes

Lierop, W.F.J.; Nijkamp, P.

1978

A UTILITY FRAMEWORK FOR INTERACTION MODELS FOR SPATIAL PROCESSES

Wal van Lierop
Peter Nijkamp

Research Memorandum 1978-12

Paper presented at the
Advanced Summer Institute
in Regional Science,
Siegen,
August 1978.

1. Introduction

In the post-war period human settlement and mobility patterns have shown rapid and significant changes. Urbanization and spatial agglomeration were a first major trend, followed by a large-scale suburbanization movement and an urban decay in a broader sense. At present a more diffuse pattern emerges: on the one hand, the suburbanization movement appears to continue as a spatial movement towards more remote rural and even peripheral areas, while on the other hand some big cities tend to start again acting as driving forces of a new agglomeration development (cf. Nijkamp [1978]).

Operational insight into the various forces determining the development of settlement and mobility is, however, rather scarce. Fortunately, during the seventies the attention of regional science, geography and transportation economics has been focussed on spatial interaction analysis in relation to the analysis of settlement and mobility patterns. A basic shortcoming, however, is the fact that the major part of these spatial interaction analyses is not based on a clear utility framework, so that the distance and attractiveness preferences of individuals and groups are hard to integrate with spatial interaction models.

This shortcoming is even more important in view of the economic trends in the western world: a reduced growth, stagnation or even decline. These trends have two aspects, viz. an economic and a spatial one. The economic changes in the western countries exert a substantial impact on production, income and investments, while the spatial changes may affect the settlement and mobility pattern (for example, due to the scarcity of energy resources). Regional growth, urban decline, energy prices, environmental constraints and global stagnation are not independent phenomena. Thus, economic and spatial changes influence each other mutually, so that it is extremely important to analyze the spatial processes in a detailed and adequate manner.

This implies that spatial mobility trends are to be studied at a micro scale, viz. (1) at the scale of individual (or disaggregated group) preferences and (2) at a detailed spatial scale. Consequently, a prerequisite for analyzing spatial impacts of stagnation is the construction of a spatially disaggregated and utility-based interaction model. The construction of these models is also important to judge the feasibility, desirability and consistency of public policies and public projects, for example in the field of industrialization, new investments, infrastructural facilities, and physical planning. These public policies and projects will only be successful, if the determinants of human spatial behaviour are sufficiently known.

In spite of an impressive box of utility theories, fairly little progress has been made to develop operational utility-based spatial interaction models which are able to overcome the disaggregation problem of preferences and priorities of individuals, households and groups.

The present paper is devoted to the methodology of disaggregated spatial decisions. Attention will be paid inter alia to modal choice models for individuals and households. A utility framework will be developed in order to analyse mobility patterns and traffic flows on the basis of modern variants of consumer theory. Special attention will be devoted to a logit approach to spatial interactions. Finally, a formal relationship will be discussed between the utility background of the logit model and the traditional entropy model.

2. Methodological Remarks

As set out above, any physical planning needs insight into the determinants of spatial behaviour. Especially in a period of stagnating growth an optimal and effective allocation of public funds in the field of infrastructural, environmental and urban planning is necessary. A prerequisite for adequate physical planning is information on present and future traffic volumes, causes of spatial mobility processes and impacts of alternative public policies.

The analysis of mobility patterns may pertain to short-run and medium- and long-run developments.

Short-run analysis aims at identifying the determinants of spatial flows and processes on the basis of observed mobility patterns and flow intensities. In this framework the effectiveness of specific mobility instruments within a global mobility and traffic policy can be studied as well. For example, Gillen [1977] has examined the effects of a differentiation in parking costs on modal choice.

Medium- and long-run analysis aims at predicting future traffic volumes for a certain area, as well as its related economic development, the spatial distribution of mobility patterns, the modal choice developments, the capacity problems of the regional network and the environmental impacts. An example of the latter type of analysis is the Integral Traffic and Transportation Study of the Dutch Ministry of Traffic and Water Management [1972] (see also Steenbrink [1974]).

It is clear that the medium- and long-run analysis is based on a more comprehensive and integrated approach, in which mobility processes are codetermined by broader economic, social and spatial trends. This approach incorporates also a larger set of structural variables which may effect the mobility variables. A necessary consequence of these medium- and long-term models is a fairly aggregate level of analysis, while the short-term models open more opportunities for incorporating qualitative aspects and more disaggregated methods.

In the past both types of analysis have been elaborated in spatial research, inter alia by means of multiple regression analysis, category analysis, gravity and entropy analysis and electro-static field analysis. The majority of these techniques, however, were hardly capable to attack effectively the urgent problems of transportation and traffic planning (see Bouchard [1974]). The most important objections against the abovementioned techniques are: the lack of theoretically-founded causal relationships, the abandonce of several relevant factors in mobility decisions and the (usually) aggregate nature of transportation and traffic models.

Several attempts have been undertaken to improve the methodological basis of the abovementioned techniques. For example, several utility foundations have been developed as a justification for the use of gravity models (cf. Isard [1960] and Nijkamp [1975]) and entropy analysis (cf. Nijkamp [1978]).

In addition to this trend, a new approach emerging from psychology has been developed, viz. the construction of disaggregated choice models. These models represent the behaviour of individuals in contrast with the previous models which focus mainly on groups of persons in a certain area. These disaggregated simultaneous choice models aim at providing descriptive and prediction models on the basis of individual behavioural relationships and underlying determinants. In this respect an attempt may be made to develop a model which is more applicable to regions and population groupes marked by comparable or similar properties.

At the beginning the disaggregated choice models were mainly employed in short-term mobility research in order to find a better explanation for traffic and transportation phenomena and to investigate the effectiveness of some instruments of mobility policy. For example, McGillivray [1972] and Warner [1962] have applied binary choice models [1)] to the San Francisco Bay and to Chicago respectively in order to examine the impacts of changes in travel time and travel costs upon modal choice.

The succes of short-term disaggregated models has stimulated the development and use of disaggregated choice models as instruments for

1) Binary choice models are choice models with only two alternatives (for example, private cars and public transport).

analyzing long-term developments in traffic and transportation. Examples can be found among others in McFadden [1975] (in a study of Pittsburgh and its surroundings), Richards and Ben-Akiva [1975] (in a study of Eindhoven), and Westin and Watson [1974,1975] (in a study of Edinburgh and Glasgow). The present paper will mainly focus on the methodology of an important class of disaggregated choice models, viz. the logit approach.

Disaggregated spatial choice models have to be based on behavioural assumptions. Consequently, the specification of such a model may benefit from the theory of consumer behaviour. The conventional consumer theory on optimal allocational decisions of a rational consumer is, however, too restrictive as a straightforward theoretical foundation of optimal spatial choice behaviour, because of its static nature, the divisibility hypothesis of continuous commodity sets, the presence of only one (budget) constraint, and the absence of distance. It is clear, however, that it is a heavy task to build a dynamic utility model of spatial choice behaviour, to incorporate mutually exclusive choice alternatives, to include additional constraints (time budget, network capacities etc.), and to take account of spatial co-ordinates in choice analysis.

In spite of those frictions it is reasonable to require that an effective micro-economic theory of consumer behaviour for trip decisions has to describe how an individual consumer takes a certain spatial decision from a finite set of mutually exclusive trip alternatives (composed of commodities/services and time which are mixed in discrete proportions for each alternative), such that his utility is at a maximum, taking into account the restrictions imposed <u>inter alia</u> by his income and time budget.

It is plausible to assume that the value of time is co-determined by the evaluation of the circumstances under which travel time is spent (for example, travel comfort). This implies that a utility maximum is characterized by a situation where the marginal utilities of all commodities are equal, in the sense that the marginal utility of time should be interpreted as its total marginal value (including the time evaluation effect). This so-called comfort effect (cf. De Donnea [1971] and Van Lierop and Veenema [1976]) means that the (first order) Slutsky conditions [1] for a utility maximum should not only relate the spatial trip choice possibilities in the consumer's consumption to his income, but also to his personal evaluation of travel time for alternative trips. In this way the micro-economic consumer theory may provide a useful basis for disaggregated

---

1) An exposition of Slutsky conditions can be found among others in Somermeyer [1967].

spatial choice models. This will be exposed in more formal terms in section 3.

3. Theoretical Formulation of the Disaggregated Choice Model

In general the specification of an economic relationship has to take into account conditions of a logical, theoretical, methodological and practical nature (see Somermeyer [1967] for a detailed explanation of this subject). In the following theoretical foundation of disaggregated choice models an attempt will be made to meet these requirements as much as possible.

Disaggregated choice models are always stochastic, so that they are subject to a probability distribution. So their ultimate choice probability falls between 0 and 1. It is, however, difficult to identify the exact probability, because we can only observe that a given person has chosen a certain alternative or not. Let us assume, that the expectation of the realization of the ultimate choice (Y) for a certain alternative i equals its unobservable probability, P_i. In other words, $E(Y_i) = P_i$, in which Y_i is a so-called Bernoulli random variable (that means it can only have a value 1 or 0), while its expectation $E(Y_i)$ and, consequently also the probability P_i, are continuous at the 0 - 1 interval. When, given the micro-economic consumer theory, a vector of explanatory variables is constructed and next a set of observations on these variables is made by means of an inquiry, the values of these variables often appear to fall outside the 0 - 1 interval, which defines the probability P_i. Consequently, it will be difficult to find a direct relationship between a chosen alternative and its individual utility function (for instance, it seems to be impossible to use regression analysis). It is, however, possible to formulate a meaningful model by applying a transformation upon the probability distribution, as is shown in Fig. 1.

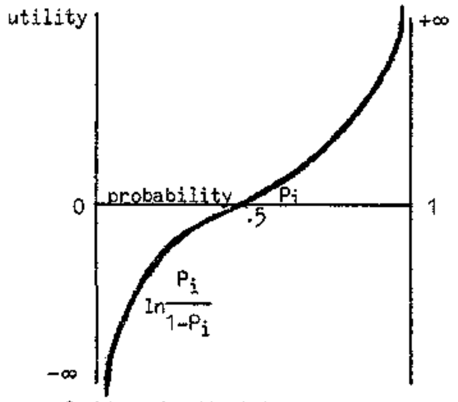

Fig. 1. Transformation of the individual utility function (defined at the range [- ∞;+ ∞]) to the probability P_i (defined at the interval [0;1]) by means of the so-called logit, specified as: $\ln[P_i/(1-P_i)]$.

The probability P is measured along the horizontal axis and its transformation along the vertical axis. When P increases from 0 to 1, its transformation varies between $-\infty$ and $+\infty$. This approach solves the interval problem.

There are a lot of transformation methods which possess the latter quality. The most important methods are: the linear probability model, discriminant, probit and logit analysis. The latter method is preferred mainly because it is relatively simple to handle and it has some other real good properties, such as a great extent of flexability concerning the specification of the different equations (utility functions of the several alternatives). Only if the substitutes are very close to each other, the specification of a logit model seems to be too restrictive [1]. In the present paper, the logit model will take an important place.

When we write the utility of trip alternative i for consumer t as U_{it}, the probability that this consumer chooses alternative i can be described as follows:

$$(3.1.) \quad \underline{P}(i:A_{Nt}) = \text{Prob}[\underline{U}_{it} \geq \underline{U}_{jt} \;,\; \forall j \in A_{Nt}]$$

in which j may be any alternative out of a finite set of (each other mutually excluding) alternatives (A_N) consumer t has at his disposal. Equation (3.1.) tells us that the probability to choose one special alternative out of a whole range ($\underline{P}(i:A_{Nt})$), equals the probability that the utility of the alternative concerned is at a maximum. This utility, $\underline{U}_{it}$, is a function of a set of variables that describe alternative i explicitly. These variables may be distinguished into vectors of specific variables (X_i), generic (or non-alternative related) variables (Y_i) [2] and socio-economic variables (S_t) characterizing consumer t. An example of a specific variable is travel time by public transport or by private car in a modal choice model. A feature of a generic variable is that it has the same influence in all alternatives it is related to. Consequently, its coefficient has the same value in all equations. The vector S_t contains variables with mixed specific-generic aspects. Elements of S_t may be income and valuation of time. These variables open a possibility to incorporate budget and time constraints in a utility function.

A simple notation of the utility arising from a choice in favour of alternative i to consumer t is:

$$(3.2.) \quad \underline{U}_{it} = u_i(X_i, Y_i, S_t) + \underline{\eta}_{it}$$

1) See Hausman and Wise [1978].
2) See Van Essen [1977].

with $\underline{n}_{it}$ is a stochastic term reflecting the casual elements of utility. Substitution of equation (3.2.) in (3.1.) gives

$$(3.3.)\quad \underline{P}(i:A_{Nt}) = \text{Prob}[u_i(X_i,Y_i,S_t) + \underline{n}_{it} \geq u_j(X_j,Y_j,S_t) + \underline{n}_{jt},\ (\forall j \varepsilon A_{Nt})]$$

which can be rewritten to

$$(3.4.)\quad \underline{P}(i:A_{Nt}) = \text{Prob}[u_i(X_i,Y_i,S_t) - u_j(X_j,Y_j,S_t) \geq \underline{n}_{jt} - \underline{n}_{it},\ (\forall j \varepsilon A_{Nt})]$$

The latter relationship states that consumer t will choose trip alternative i when the difference between those parts of the utility of the alternatives i and j, which represent the various variables, is greater than or equal to the difference between their stochastic components.

In general, it is assumed this function (3.4.) will be S-shaped [1]. In that case, however, the abovementioned linear probability method is less useful. The probability that a certain trip alternative will be chosen is in that case very high or very low; in other words, the probability will have a value near 1 or 0. Intuitively, this can be explained by the seemingly strong psychological and social rigidities in attitudes of many people, with regard to certain aspects of a given trip alternative, so that they will almost never change their attitude even when the relative benefits of other alternatives would show a sharp increase.

If we assume now that the random disturbances of the various alternatives are independent and equally Weibull distributed[2], these terms cancel each other out in a cumulative distribution function (G) of the difference between both. This means that equation (3.4.) can be formulated as [3]:

$$(3.5.)\quad \underline{P}(i:A_{Nt}) \equiv [G(u_i(X_i,Y_i,S_t) - u_j(X_j,Y_j,S_t)),\forall j \varepsilon A_{Nt}] = \frac{e^{u_i(X_i,Y_i,S_t)}}{\sum_{j=1}^{A_{Nt}} e^{u_j(X_j,Y_j,S_t)}}$$

In formal terms, this is the general form of a stochastic simultaneous, multinomial logit model. The specification to be chosen for the exponent $[u_j(X_j,Y_j,S_t)\ ,\ \forall j \varepsilon A_{Nt}]$ is not quite clear, however. Nevertheless, we may assume that there are no <u>a priori</u> reasons why it should not be linear [4]. We can now formulate equation (3.5.) as:

---

1) See among others Tobin [1975] or Theil [1967].
2) A Weibull distribution has the following form: $\text{Prob}(\underline{n}_j \leq n) = e^{-e^{-\underline{n}}}, (\forall j \varepsilon A_{Nt})$.
3) See Domencich and McFadden [1975] pages 53-65.
4) See for instance De Donnea [1971] chapter 3.

$$(3.6.)\quad \underline{P}(i:A_{Nt}) = \frac{e^{[\alpha_i + \sum_{k=1}^{Ki} \beta_{ik}X_{ik} + \sum_{l=1}^{L} \gamma_l Y_l + \sum_{m=1}^{Mi} \delta_{im} S_{tim}]}}{\sum_{j=1}^{A_{Nt}} e^{[\alpha_j + \sum_{k=1}^{Kj} \beta_{jk}X_{jk} + \sum_{l=1}^{L} \gamma_l Y_l + \sum_{m=1}^{Mj} \delta_{jm} S_{tjm}]}}, (t=1,\ldots,T_h)$$

with:

α_i = the constant term of alternative i (if available)

Ki = the number of specific coefficients of alternative i

β_{ik} = the k-th specific coefficient of alternative i

X_{ik} = the k-th specific variable of alternative i

L = the number of generic coeffients

γ_l = the l-th generic coefficient

Y_l = the l-th generic variable

Mi = the number of socio-economic coefficients for consumer t in alternative i

δ_{im} = the m-th socio-economic coefficient for consumer t in alternative i

S_{tim} = the m-th socio-economic variable for consumer t in alternative i

T_h = the group of persons - from a sample including a total number of T - to whom alternative i is a real possibility. For all remaining $(T - T_h)$ persons, the abovementioned probability equals zero, of course.

4. Problems in Specifying a General Logit Model

When we want to describe a spatial interaction system as good as possible, it is useful not to start immediately with the above formulated logit model (3.6.), but, instead, to construct some submodels. Because the nature of the demand for trips is very diverse, it is reasonable to distinguish special models for at least

a. job commuting
b. recreation
c. social and
d. business trips.

In addition it is advisable to pay some attention to the geographical origins and destinations of trips. It is clear, for instance, that the demand for trips will be influenced in another way by the vector of explanatory variables in the city than in its surroundings. Frequently even some specific variables are present at the different relations. Concerning this question Watson and Westin [1975] proved that significant differences did exist with respect to the demand for trips within and between innercities, suburbs and peripheral areas. It should be noticed, that the creation of different zones may give rise to spatial assignment problems. These will not be described here, but discussions on this subject can be found among others in Cliff et al. [1975].

Combination of the cases distinguished above shows that, when we aim to give a general and good description of the demand for trips, we have to create quite a lot of submodels, which have to be estimated separately in a consistent manner. Next we may compare the achieved results in order to discover some submodels which may be substituted for each other [1], so that the submodels concerned may be assumed to have a more general character. For example, it might be possible to investigate whether the estimators of an urban job commuting trip model can be used to forecast the commuting flows between the city and its surroundings. Should some submodels lead to more general results, then it will be meaningful to use this information as a feedback to the stage of model construction, in order to design new models for trip choices, which are suitable as forecasting models in multiple situations.

5. Estimating the Logit Model

Several methods can be used to estimate the coefficients α, β, γ and δ from equation (3.6.) with the aid of a sample of individual values for the incorporated variables. A commonly used and, from a theoretical point of view, a rather satisfactory method is, without any doubt, the method of maximum likelihood.

In general, the likelihood function of a disaggregated sample can be written as follows:

$$(5.1.) \quad \Lambda = \prod_{t=1}^{T} . \prod_{i \varepsilon A_{Nt}} \underline{P}(i:A_{Nt})^{g_{it}}$$

with:

T = the number of observations

g_{it} = 1 if person t has chosen alternative i
= 0 if not.

The log-likelihood function is:

$$(5.2.) \quad \ln \Lambda = \Lambda^* = \sum_{t=1}^{T} \cdot \sum_{i \varepsilon A_{Nt}} g_{it} \ln \underline{P}(i:A_{Nt})$$

In order to form the special likelihood function of the logit model

---

1) For tests concerning this subject, see Watson and Westin [1975].

of equation (3.6.), we have to arrange the observations in groups per alternative, in such a way that group 1 contains all T_1 individuals who actually made a choice for alternative 1 and so forth, until group n with T_n individuals who have chosen the last alternative. The likelihood functions will than be:

$$(5.3.)\quad \Lambda = \Lambda(\alpha,\beta,\gamma,\delta) = \frac{T!}{T_1!T_2!T_3!\ldots.T_n!}.\prod_{t=1}^{T_1} \underline{P}(1:A_{Nt})\ldots.\prod_{t=T_1+..T_{n-1}+1}^{T_1+\ldots T_n} .\underline{P}(n:A_{Nt})$$

and its log-likelihood specification:

$$(5.4.)\quad \ln\Lambda(\alpha,\beta,\gamma,\delta) = \hat{\Lambda} = \ln T! - \ln(T_1!T_2!\ldots.T_n!) + \sum_{t=1}^{T_1} \ln\underline{P}(i:A_{Nt}) + \ldots.$$

$$\ldots + \sum_{t=T_1+..T_{n-1}+1}^{T_1+\ldots T_n} \ln\underline{P}(n:A_{Nt})$$

After having substituted equation (3.6.) for $\underline{P}(i:A_{Nt})$ here, this function $\hat{\Lambda}$ should be maximized with respect to α_i, β_{ik}, γ_l and δ_{tim} $[(i\varepsilon A_{Nt});(k=1..K);(l=1..L);(m=1..M)]$.

There exists a maximum when the first-order derivatives of all these coefficients are equal to zero and the matrix of second-order derivatives is negative semi-definite at the same time. McFadden [1974] proved that this will always happen when the first-order conditions are satisfied [1]. In other words, the maximum of the log-likelihood function employed here is unique. Besides, the estimators of the parameters $(\alpha,\beta,\gamma,\delta)$ are maximum-likelihood estimators.

The abovementioned necessary and sufficient conditions for a maximum can be written as:

$$(5.5.)\quad \frac{\partial\hat{\Lambda}}{\partial Z_u} = 0$$

and

$$(5.6.)\quad \frac{\partial^2\hat{\Lambda}}{\partial Z_u \partial Z_u'} < 0$$

Z_u may represent any element from the vector $\vec{Z}$ which contains the parameters we have to estimate (α, β, γ and δ). Clearly, we have to

---

1) Unless the data show some very remarkable conditions; see McFadden [1973].

specify (1 + K + L + M) equations of type (5.5.) to assess all coefficients. The non-linear character of the function $\hat{\Lambda}$ causes the need of an iterative solution method for this estimation [1]. In this respect a fairly popular and practicable method is the method of Newton-Raphson [2], which has demonstrated in the past its power [3]. This method starts with (sometimes random) initial values for the vector $\vec{Z}$ of the parameters to be estimated. These initial values, denoted as $\vec{Z}_0$, are used to linearize the non-linear equation of $\hat{\Lambda}$ by means of a Taylor's series approximation. When one truncates this Taylor's series of $\partial\hat{\Lambda}/\partial\vec{Z}$ after the second term, the first necessary condition for a maximum will be:

$$(5.7.)\quad \partial\hat{\Lambda}/\partial\vec{Z}_0 = -\frac{\partial^2\hat{\Lambda}}{\partial\vec{Z}_0\partial\vec{Z}_0'}.\Delta\vec{Z}_0$$

with:

$$(5.8.)\quad \Delta\vec{Z}_0 = \vec{Z}_1 - \vec{Z}_0$$

$\vec{Z}_1$ can now be written as:

$$(5.9.)\quad \vec{Z}_1 = \vec{Z}_0 - \left[\frac{\partial^2\hat{\Lambda}}{\partial\vec{Z}_0\partial\vec{Z}_0'}\right]^{-1}.\frac{\partial\hat{\Lambda}}{\partial\vec{Z}_0}.$$

Next, in a continuous iteration process we have to determine:

$$(5.10.)\quad \vec{Z}_2 = \vec{Z}_1 - \left[\frac{\partial^2\hat{\Lambda}}{\partial\vec{Z}_1\partial\vec{Z}_1'}\right]^{-1}.\frac{\partial\hat{\Lambda}}{\partial\vec{Z}_1}$$

$\vdots$

etc.

The process stops when:

$$(5.11.)\quad \vec{Z}_w - \vec{Z}_{w-1} < \xi$$

with:

w = the number of iteration steps (o,...,w) (sometimes restricted for reasons of computer time)

ξ = a prespecified constant related to the desired precision of the approximation

$\vec{Z}_w$ = the vector with calculated values of the parameters.

---

1) For a more detailed description of some iterative solution methods see among others: Murray [1972] and Batty [1976].
2) See also Van Est and Van Setten [1977], pp. 42 - 44.
3) See among others Richards and Ben-Akiva [1975].

It is possible to prove that $(\partial^2\hat{\Lambda}/\partial\vec{Z}_w\partial\vec{Z}_w')^{-1}$ is indeed a semi-definite matrix [1].

6. Weighted Logit Analysis

When a certain alternative is underrepresented in a sample, an adjustment should be carried out. One of the simplest corrections is to add a constant term to the utility function (3.6.) of that alternative, or to raise the constant (when it is already present). This additional constant term should be large enough to guarantee that the mean probability of the alternative equals exactly the share of that alternative in the population from which the sample was drawn.

After such a correction, the model can again be estimated in the way described in chapter 5. In such a case it is called a weighted logit model [2].

7. Forecasting via Aggregation

It is useful for infrastructural policy to identify, by means of the estimated results for the individual trip choices, the total modal split in the future or in other regions. Such a forecast needs a shift from the disaggregated level of estimation to a more aggregated level. Then several difficulties will emerge mainly characterized by:

a. economic-behavioural and related aspects

b. mathematical-statistical aspects.

ad a. These problems arise due to the mutual influences of all trip-makers (e.g., congestion). Consequences may be: insufficient capacity of networks such as, for example, a shortage of parking places, traffic congestion etc. A solution to this problem may be found by including some resistance factors or shadowprices in the individual utility function of special alternatives. By doing so we could still satisfy the condition of inner consistency of an aggregated model - as formulated amongst others by Somermeyer [1967]. This condition states: the sum of the parts of an (additively) fully disaggregated total should represent that total exactly; in other words, it should be possible to construct a macro-model for any group by aggregating all micro elements belonging to that group.

ad b. A first aggregation possibility is to take together all individual probabilities at a certain alternative. Then we can calculate

---

1) For statistical tests and elasticities concerning the logit model, see: Van Lierop [1978].

2) See, for instance, Van Essen [1977].

the mean probability, $\bar{\underline{P}}_i$, for trip alternative i as:

$$(7.1.)\qquad \bar{\underline{P}}_i = \frac{\sum_{t=1}^{T} \bar{\underline{P}}(i:A_{Nt})}{T} \qquad 1)$$

A rough measure for the total demand for trip alternative i would thus be: $T.\bar{\underline{P}}_i$.

For the short run, a much better forecast can be made by means of the aggregated elasticity, i.e., the elasticity with respect to the abovementioned mean probability. Let us assume:

$$(7.2.)\qquad \frac{dV_{ti\theta}}{dV_{i\theta}} = \frac{V_{ti\theta}}{V_{i\theta}} \quad \text{for } t = 1,2,\ldots\ldots,T$$

In which $V_{ti\theta}$ may be any element of the vector of θ(=K + L + M) variables X, Y and S, which are included as arguments in the utility function of alternative i for consumer t (see equation (3.6.)), and:

$$(7.3.)\qquad V_{i\theta} = \frac{\sum_{t=1}^{T} V_{ti\theta}}{T}$$

So the ratio of the change in any variable θ of alternative i for consumer t to the mean change of the same variable of alternative i equals the ratio of the total influence of that variable of alternative i for consumer t to the mean total influence of θ in alternative i. In other words, expression (7.2.) means that the proportional change of $V_{i\theta}$ is the same for every person. It is possible to formulate the aggregated elasticity now as:

$$(7.4.)\qquad \xi_{V_{i\theta}}^{\underline{P}_i} = \frac{\dfrac{d\bar{\underline{P}}_i}{dV_{i\theta}}}{\dfrac{\bar{\underline{P}}_i}{V_{i\theta}}} = \frac{\sum_{t=1}^{T} \underline{P}(i:A_{Nt}).\xi_{V_{ti\theta}}^{\underline{P}(i:A_{Nt})}}{\sum_{t=1}^{T} \underline{P}(i:A_{Nt})}$$

in which $\xi_{V_{ti\theta}}^{\underline{P}(i:A_{Nt})}$ is the so-called direct elasticity of the choice of a certain individual t for alternative i with respect to variable θ. The numerator of the right part of equation (7.4.)

1) It is assumed that the probabilities in this sum are independent.

gives a reasonable good short run guess of the change of the total demand for trip alternative i. It may be necessary to add a correction factor $V_{i\theta}/V_{ti\theta}$ to this numerator because the initial appreciation of a certain variable θ may not be the same for everybody. Finally, it might be remarked that the aggregate elasticity is in fact a weighted sum of the individual elasticities, with as weighting factors the individual probabilities (see also Richards and Ben-Akiva [1975]).

A method which is more complete in many respects is the construction of the relative frequency distribution of the probabilities for a trip alternative i for the population, given the information on the relative frequency distribution of V_i (the matrix of variables that characterize alternative i). First we rewrite our logit model, analogous to Westin [1973], in a slightly different form [1]:

$$(7.5.)\quad \ln\left\{\frac{\underline{P}(i:A_{Nt})}{1-\underline{P}(i:A_{Nt})}\right\} = -\ln \sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j - \vec{Z}'_{iw}\vec{V}_i)} \qquad 2)$$

---

1) Westin uses: $\ln\frac{P_i}{1-P_i} = X_i\beta$

2) The derivation can be made as follows:

$$\underline{P}(i:A_{Nt}) = \frac{e^{\vec{Z}'_{iw}\vec{V}_i}}{\sum_{j=1}^{N} e^{\vec{Z}'_{jw}\vec{V}_j}} = \frac{1}{\sum_{j=1}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}} = \frac{1}{1+\sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}}$$

and:

$$\left\{1-\underline{P}(i:A_{Nt})\right\} = \frac{1+\sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}}{1+\sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}} - \frac{1}{1+\sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}}$$

$$= \frac{\sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}}{1+\sum_{\substack{j=1\\ j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}}$$

(to be continued at page 15)

The construction of the distribution for the population $(\underline{P}(i:A_N))$, needs two transformations, viz. one from $\vec{V}_i$ to $\vec{Z}'_{iw}\vec{V}_i$, and the other one from $\vec{Z}'_{iw}\vec{V}_i$ to $\underline{P}(i:A_N)$; $(\forall i \in 1,\ldots,N)$. In case the distribution of $\vec{Z}'_{iw}\vec{V}_i$ is known, it will be possible to form its probability density function $g(\vec{Z}'_{iw}\vec{V}_i)$. When there are just 2 alternatives, then:

$$(7.6.)\qquad \ln\left\{\frac{\underline{P}(i:A_N)}{1-\underline{P}(i:A_N)}\right\} = \vec{Z}'_{1w}\vec{V}_1 - \vec{Z}'_{2w}\vec{V}_2$$

will be the logit formula. Next, one may specify by means of $g(\vec{Z}'_{iw}\vec{V}_i)$ the density function of the probabilities of alternative i, viz. $f\{\underline{P}(i:A_N)\}$. Assume that the distribution of $\vec{Z}'_{iw}\vec{V}_i$ will be continuous and that $\vec{V}_i$ will have a multidimensional normal distribution with a row vector of means μ_V and covariation matrix Σ. Then $\vec{Z}'_{iw}\vec{V}_i$ will have a onedimensional normal distribution with mean $\mu = \vec{\mu}_V\vec{Z}_{iw}$, and variation $\sigma^2 = \vec{Z}'_{iw}\Sigma\vec{Z}_{iw}$. Then it is possible to rewrite the probability function in the two-alternative case from

$$(7.7.)\qquad f\{\underline{P}(i:A_N)\} = g\left(\ln\left\{\frac{\underline{P}(i:A_N)}{1-\underline{P}(i:A_N)}\right\}\right)\cdot\frac{1}{\underline{P}(i:A_N)\{1-\underline{P}(i:A_N)\}}$$

to the normal distributed form:

$$(7.8.)\qquad f\{\underline{P}(i:A_N)\} = \frac{1}{\sqrt{2\Pi\sigma^2}}\cdot\frac{1}{\underline{P}(i:A_N)\{1-\underline{P}(i:A_N)\}}\cdot\exp\left(-\frac{1}{2\sigma^2}\left\{\ln\frac{\underline{P}(i:A_N)}{1-\underline{P}(i:A_N)}-\mu\right\}^2\right)$$

This is the form of a so-called S_B probability density function as defined by Johnson [1949].

Direct estimation of $\underline{P}(i:A_N)$ is not possible. That is the reason

---

2) of page 14 continued:
Combination now gives:

$$\frac{\underline{P}(i:A_{Nt})}{1-\underline{P}(i:A_{Nt})} = \left\{\sum_{\substack{j=1\\j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}\right\}^{-1}$$

This leads to:

$$\ln\left\{\frac{\underline{P}(i:A_{Nt})}{1-\underline{P}(i:A_{Nt})}\right\} = -\ln\sum_{\substack{j=1\\j\neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j-\vec{Z}'_{iw}\vec{V}_i)}$$

The left-hand part of the latter formula is called the logit.

why an iterative procedure should be used. First, consistent estimators for the means $\vec{\mu}_V$ and variance Σ of the explanatory variables $\vec{V}_i$ are taken from the sample. The values of the coefficients $\vec{Z}_i$ are also known from the estimation of equation (3.6.). Thus it is possible to estimate μ and σ^2. As set out above, the μ and σ^2 are linear combinations of the mean and the (co-)variances of $\vec{V}_i$, respectively. Therefore, the effect of each change on one of the values of $\vec{V}_i$ may also be expressed directly in terms of the change of the first and second moment of the probability function. Should just any variable $V_{i\theta}$ increase with a constant term c to get the new value $V_{i\theta}$ + c, then the S_B-function of $\underline{P}(i:A_N)$ may be characterized as:

(7.9.) $(\mu + Z_{i\theta}.c, \sigma^2)$.

When the change has a multiplicative nature, say $(1 + c).V_{i\theta}$, the parameters of the S_B-function become:

(7.10.) $(\mu + cZ_{i\theta}\mu_\theta, \sigma^2 + Z_{i\theta}(2c + c^2)\sigma^2_\theta + 2 \sum_{\zeta\neq\theta} Z_{i\zeta}Z_{i\theta}c\sigma_{\zeta\theta})$

in which μ_θ, σ^2_θ and $\sigma_{\zeta\theta}$ are the mean, the variance and the covariance of the corresponding components of $\vec{V}_i$, respectively.

By means of the relative frequency distribution (7.8.) the expectation of the part of that population that chooses a certain alternative can be determined. This can be defined as follows:

(7.11.) $E[\underline{P}(i:A_N)] = \int_0^1 \underline{P}(i:A_N) f\{\underline{P}(i:A_N)\} d\{\underline{P}(i:A_N)\}$

When some change takes place in the vector of variables $\vec{V}_i$ which characterize the choices, it will be interesting to know how this effects $E[\underline{P}(i:A_N)]$. A small change may be analysed by calculating the elasticities of $E[\underline{P}(i:A_N)]$, in relation with the moments of $\vec{V}_i$. In case of bigger fluctuations it will however be necessary to adjust the transformations of the relevant explanatory variables and their coefficients to the distribution function of the population.

The benefit of employing a frequency distribution for the probabilities of the population compared with the common use of an extrapolation of the probability of the mean individual as a forecasting measure is, that the first one will likely supply better results in the short run [1]. For long-run purposes it is highly questionable whether the abovementioned difference should be taken into account. It is rather likely that then the use of disaggregated choice models will become questionable. Especially as these models have no time dimension at all, they can hardly be suitable to forecast the far remote future. For example, then it will not be possible to discover a trend in the demand for trip alternatives by which the need for subsequent years could be predicted [2].

---

1) This may be explained as follows: let

$$\xi_{V_{im}} = \frac{Z_{i\theta}\mu_\theta . E[\underline{P}(i:A_N)]\{1 - E[\underline{P}(i:A_N)]\}}{E[\underline{P}(i:A_N)]}$$

be the elasticity of the mean individual m with respect to a change of those variables determining his choice for a certain alternative i. The relation between this aggregated elasticity and the one computed from the frequency distribution of the population, which can be described by:

$$\xi_{V_i} = \frac{Z_{i\theta}\mu_\theta . E[\underline{P}(i:A_N)\{1 - \underline{P}(i:A_N)\}]}{E[\underline{P}(i:A_N)]} ,$$

may be expressed as:

$$\frac{E[\underline{P}(i:A_N)]\{1 - E[\underline{P}(i:A_N)]\}}{E[\underline{P}(i:A_N)\{1 - \underline{P}(i:A_N)\}]}$$

This can be rewritten as:

$$\frac{E[\underline{P}(i:A_N)] - \{E[\underline{P}(i:A_N)]\}^2}{E[\underline{P}(i:A_N)] - E[\underline{P}(i:A_N)]^2}$$

Given the fact that:

$$\text{var } \underline{P}(i:A_N) = E[\underline{P}(i:A_N)]^2 - \{E[\underline{P}(i:A_N)]\}^2 \geq 0 ,$$

we may conclude:

$$E[\underline{P}(i:A_N)]^2 \geq \{E[\underline{P}(i:A_N)]\}^2 .$$

This means: the abovementioned quotient of the elasticities is greater than or equal to 1. So the elasticities based on the mean individual tend to overestimate the real aggregated elasticity.

2) For objections against the use of long-term forecasts with logit models, see also Ruijgrok [1978].

8. Problems in Forming the Distribution Function

In the preceding section the probability density function has been derived by using only 2 alternatives in equation (7.5.). When this assumption is dropped, (that means: each individual may make a choice out of more than 2 alternatives), this will lead to serious difficulties in equation (7.5.). This equation has to deal with the following formation:

$$\sum_{\substack{j=1 \\ j \neq i}}^{N} e^{(\vec{Z}'_{jw}\vec{V}_j - \vec{Z}_{iw}\vec{V}_i)}$$

in which the separate N elements have a lognormal distribution. The aggregation of lognormal distributions is a problem not yet solved in statistics. So in the case of more than two alternatives, there is no solution for the moment.

A reverse procedure, by forming the distribution by means of the moment generating function is not possible, because this function is not useful in the lognormal case [1).].

Such a method might be helpful however, if one assumes another distribution for the explanatory variables $\vec{V}_i$. For instance, a Poisson distribution, analogous to Hamerslag [1977], might be useful [2)]. The elements characterizing the several alternatives, however, are often distributed continuously. It is reasonable then to treat different variables, with a 0 - 1 nature as dummies. The only effect will be that, if they have a value 1, they cause a shift in the original logit formula the size of the value of their coefficients [3)]. The choice of a Poisson distribution could only be justified, when the vector of explanatory variables consists almost exclusively of discrete distributions.

Another (theoretical) solution might be to assume independency among the alternatives by defining the possibilities for each of them (either chosen or not) by the quotient of the probability at any alternative with respect to one minus that probability. In our opinion, however, it is unrealistic to suppose that the choice of a specific trip alternative would not influence other trip possibilities. So the latter method seems to be a less meaningful approach.

Obviously, a kind of "pat"-position is reached with the logit model for more than 2 alternatives. Therefore, the design of justifiable spatial

---

1) See Mood, Graybill and Boes (1974) page 541.
2) His startingpoint, however, is somewhat different.
3) An additional difficulty emerges when the dummies are dependent on each other or on normal explanatory variables. In such a case adjusted transformations should be applied. See Westin [1977].

interaction models for more than two alternatives with the aid of the logit model is fraught with difficulties.

9. Solutions for Some Problems of the Logit Model

In this section, attention will be paid to solve some problems inherent in the use of the logit model. For example, Hausman and Wise [1978] were able to attack some of the difficulties of the logit approach by constructing a "conditional probit model". A main characteristic of this model is the possibility of correlation among the random disturbances in the individual utility function. Consequently it is not necessary to assume independency among the alternatives, like in the logit model. The disaggregated choice models estimated by means of their method tend to provide better estimation results than the more traditional probit models, so that we may tentatively conclude, that they have developed a more suitable approach for analyzing the individual trip behaviour.

Other possibilities to solve some severe logit problems might be offered by the entropy model. Hereafter we will pay some attention to the relationships and similarities between entropy models and logit models. The starting point of the analysis is the entropy concept from the information theory:

$$(9.1.)\quad H = \underline{P}(i:A_N)h[\underline{P}(i:A_N)] + \{1 - \underline{P}(i:A_N)\}h\{1 - \underline{P}(i:A_N)\}$$

$$= \underline{P}(i:A_N)\ln\frac{1}{\underline{P}(i:A_N)} + \{1 - \underline{P}(i:A_N)\}\ln\frac{1}{\{1 - \underline{P}(i:A_N)\}}$$

with: H = the entropy

$\underline{P}(i:A_N)$ = the probability that alternative i will be chosen from a set of N available alternatives

$h[\underline{P}(i:A_N)] = \ln[1/\underline{P}(i:A_N)] = -\ln\underline{P}(i:A_N)$ = the decreasing function on which the information approach is based (see Theil [1972]).

Next, we will write the logit, L, as:

$$(9.2.)\quad L = \ln\left(\frac{\underline{P}(i:A_N)}{1 - \underline{P}(i:A_N)}\right)$$

(see also note 2 of page 14)

When we consider the effect of a change in the probability $\underline{P}(i:A_N)$ on the entropy (in other words, the influence of the marginal probability on the entropy), we get:

$$(9.3.)\quad \frac{dH}{d[\underline{P}(i:A_N)]} = \frac{d}{d[\underline{P}(i:A_N)]}.[-\underline{P}(i:A_N)\ln\underline{P}(i:A_N)-\{1-\underline{P}(i:A_N)\}\ln\{1-\underline{P}(i:A_N)\}]=$$

$$= - \ln\{\underline{P}(i:A_N)\} - 1 + \ln\{1 - \underline{P}(i:A_N)\} + 1 =$$

$$= - \ln\left\{\frac{\underline{P}(i:A_N)}{1 - \underline{P}(i:A_N)}\right\} = - L$$

This implies that the derivative of the entropy to the probability is: minus the logit (see Theil [1972]). We may interpret this result in the sense that the logit measures the sensitivity of uncertainty for variations in the probability.

We can also view this problem from the reverse side by rewriting (9.3.) as:

$$(9.4.) \quad \int \frac{dH}{d[\underline{P}(i:A_N)]} . d[\underline{P}(i:A_N)] = -\int L \, d[\underline{P}(i:A_N)] \quad .$$

This leads to:

$$(9.5.) \quad \int L d[\underline{P}(i:A_N)] = - H \quad .$$

In other words, a change in the probability of a certain alternative, will change the entropy also. Intuitively, this can be illustrated as follows: when the choice of a certain trip alternative is marginal (so the probability that it will be realized is very small), then the message stating this alternative will nevertheless be chosen contains a great deal of information. Now, when the choice becomes less unique (that is, when the probability increases), the amount of additional conveyed information will decrease, so that also the entropy decreases in value. This means, in terms of the abovementioned theories (see equation (9.3.)): when the probability increases, the logit increases proportionately with a simultaneous decrease of the marginal entropy (or inversely). We may clarify this result by Fig. 2.

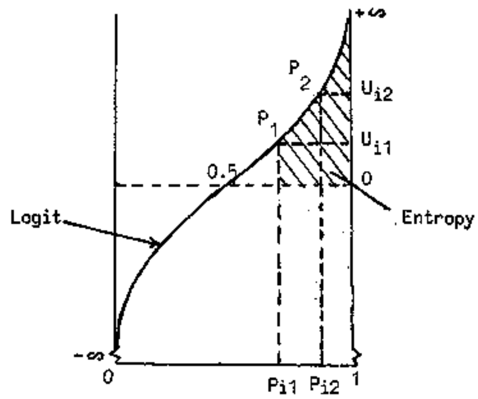

Fig. 2. Graphical description of the relationship between entropy and logit. Here, u_{i1}, is the expression of the value (or utility) of the vector of variables characterizing a given alternative i, with respect to the choice

probability P_{i1}. The surface below the logit function L represents the total entropy. Only the positive part is of current interest. So the initial entropy is represented by the shaded area. Suppose, the probability changes; for example, let P_{i1} shift to P_{i2}. Then, the probability of alternative i increases, while the entropy decreases with the shaded (positive) part of area $P_{i2}P_2P_1P_{i1}$.

One may question wether the relation between logit and entropy as described here will help us in attacking some frictions of the logit approach. One of the aims of this paper was to use the entropy model as a possibility to avoid the problems we were confronted with in the logit model. Concerning this specific question, fairly little progress has been made so far. In the current logit approach, the independency problem still remains a serious problem. Because it is impossible to handle simultaneously probabilities for multiple choice alternatives in a logit model, we are still confronted with a binary situation, in which we implicitly assume independency among the various alternatives.

On the other hand, by considering an aggregated logit model in terms of the same relations as an entropy model, we have found a real, clear and disaggregated utility foundation for the entropy model. Consequently, the important result is derived that the micro-economic consumer theory for trip choices may serve not only as a methodological basis for the logit model, but may also be associated indirectly with the entropy model via a marginality interpretation.

10. Conclusions

The construction of spatial interaction models based on individual utility theories is a process with many difficulties. Clearly, the development of logit models and the discovery of an micro-utility foundation for the entropy model involved some progress, but on the other hand, for the time being, the logit models are still unsatisfactory when the assumption of a binary situation is dropped.

In case of more than 2 trip alternatives, it is impossible to derive satisfactory estimators for the various moments of the probability density function of trip-makers, so that a forecasting model is hard to derive. Consequently, when estimators are still presented in such situations, the independency is invalidated. In this respect, only the conditional probit model of Hausman and Wise [1978] may offer some better perspectives.

Another future research problem is the inclusion of dynamic elements in logit, probit and entropy models, so that the potential of these models to forecast long-run trends can be improved. In this way, disaggregated choice versions of the so-called STARIMAR-model [1], originally formulated among others by Cliff et al. [1975], may be useful.

---

1) STARIMAR means: Space-time autoregressive integrated moving average with additional regression terms.

References

Batty, M., Urban Modelling, Algorithms, Calibrations, Predictions, Cambridge University Press, 1976.

Becker, G., A Theory of the Allocation of Time, The Economic Journal, September 1965, pp. 493-517.

Beckmann, M.J., Entropy, Gravity and Utility in Transportation Modelling, Information, Inference and Decision (ed. Menges, G), Reidel, Dordrecht, 1974

Ben-Akiva, m., and Atherton, T.J., Methodology for Short-range Travel Demand Predictions (analysis of carpooling incentives), Journal of Transport Economics and Policy, Sept. 1977, pp. 224-261.

Bouchard, R.J., Relevance of Planning Techniques to Decision Making, HRB special report no. 143, Highway Research Board, Washington DC, 1974.

Le Clercq, F., Een Amsterdams Verkeers- en Vervoersmodel: Combinatie van Aggregate en Disaggregate Technieken. Colloqium Vervoersplanologisch Speurwerk 1977.

Cliff, A.D., Haggett, P., Ord, J.K., Bassett, K.A., and Davies, R.B., Elements of Spatial Structure, a quantitative approach, Cambridge University Press, 1975.

De Donnea, F.X., The Determinants of Transport Mode Choice in Dutch Cities, Universitaire Pers Rotterdam, 1971.

Domencich, T.A., and McFadden, D., Urban Travel Demand, a behavioral analysis, North-Holland Publishing Company, Amsterdam,1975.

Van Essen, P.G., Schattings- en Calibratieprogramma's Vervoersmodellen, IWIS-TNO Den Haag, juni 1977.

Van Est, J., en Van Setten, J., Calibration Methods: Application for some spatial distribution models, Planologisch Studiecentrum TNO, Delft, April 1977.

Gillen, D.W., Estimation and Specification of the Effects of Parking Costs on Urban Transport Mode Choice, Urban Economics, May 1977.

Hamerslag, R., Onderlinge Afhankelijkheid van Vervoervraag, Planologische Ontwikkelingen en Verkeersplanning, Verkeerskunde, september 1975.

Hamerslag, R., Verkeerskundige modellen II; Deel E: Multiproportionele Poisson Schattingsmodellen- en Toetsingsmodellen, D.H.V. Amersfoort, Assen, Hengelo, 1977.

Hausman, J.A., and Wise, D.A., A Conditional Probit Model for Qualitative Choice: Discrete Recognizing Interdependence and Heterogeneous Preferences, Econometrica, Vol.46, No.2, March, 1978.

Isard, W., Methods of Regional Analysis, M.I.T. Press, Cambridge, Mass., 1960.

Johnson, N.L., Systems of Frequency Curves Generated by Methods of Translation, Biometrika 36, 1949, pp. 149-176.

Keeping, E.S., Statistical Interference, D. v. Nostrand Company, Inc. Princeton, New Yersey, 1962, pp. 123 and further.

Kraft, A., and Kraft, J., Preference Orderings as Determinants of Transport Mode Choice, Regional Science and Urban Economics 5, 1975, pp. 251-261.

Kraft, J., and Kraft, A., Emperical Estimation of the Value of Travel Time Using Multi Mode Choice Models, Journal of Econometrics 2, 1974, pp. 317-326.

Van Lierop, W., en Veenema, F., Het Micro-Economische Consumentengedrag als Theoretische Fundering van de Modal Choice, Werkstuk vervoerseconomie, V.U. 1976.

Van Lierop, W., Enige Beschouwingen over het Meten van de Vraag naar Verplaatsingen in het Personenvervoer, in: Vervoersekonomische Kaleidoskoop, bijdragen aangeboden aan Prof.Dr. J.P.B. Tissot van Patot onder redaktie van Drs. H.A. van Gent en Prof.Dr.P.Nijkamp, V.U. - Amsterdam, februari 1978.

Van Lierop, W., Micro-Economisch Consumentengedrag en Personenvervoer, Doctoraalscriptie Regionale Economie V.U. - Amsterdam, 1978.

McCarthy, P.S., Residential Location and the Journey to Work (an empirical analysis), Journal of Transport Economics and Policy, May 1977.

McFadden, D., Conditional Logit Analysis of Qualitative Choice Behavior, Frontiers in Econometrics (P. Zarembka ed.), Academic Press, New York, 1973.

McFadden, D., The Measurement of Urban Travel Demand, Journal of Public Economics 3, 1974, pp. 303-328.

McGillivray, R., Binary Choice of Urban Transport Mode in the San Francisco Bay Region, Econometrica 40 no. 5, Sept. 1972, pp. 827-848.

Morrison, D.G. Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions, Journal of the American Statistical Association, March 1972, Volume 67, Number 337.

Mood, A.M., Graybill, F.A., and Boes, D.C., Introduction to the theory of Statistics, Third Edition, McGraw-Hill, 1974.

Murray, W., Numerical Methods for Unconstrained Optimization, Academic Press, London and New York, 1972.

N.E.I., Integrale Verkeers- en Vervoersstudie, annex III, Personenvervoer, Staatsuitgeverij, Den Haag, 1972.

Neter, J., and Maynes, E.S., On the Appropriateness of the Correlation Coefficient with a 0,1 Dependent Variable, Journal of the American Statistical Association, June 1970, Vol.65, Number 330.

Nijkamp. P., Reflections on Gravity and Entropy Models, Regional Science and Urban Economics, vol. 5, 1975, pp. 203-225.

Nijkamp, P., Gravity and Entropy Models: The State of the Art, Research-Memorandum no. 1977-2, V.U. Amsterdam.

Nijkamp, P., A Spatial Complex Analysis of Agglomeration and Settlement Patterns, Research Memorandum RM-78-9, International Institute for Applied Systems Analysis, Laxenburg, 1978.

Pindyck, R.S., and Rubinfeld, D.L.,Econometric Models and Economic Forecasts, New York, McGraw-Hill, 1976.

Project Bureau IVVS, Ministry of Transport the Netherlands, SIGMO Study, part 1, Introduction and Summary, 1977.

Richards, M.G., and Ben-Akiva, M.E., A Disaggregate Travel Demand Model, Saxon-House, Farnborough, 1975.

Ruijgrok, C.J., Het Logit Model - een strak keurslijf of niet? Colloqium Vervoersplanologisch Speurwerk, 1976.

Ruijgrok, C.J., De Toepassing van een Autobeschikbaarheidsmodel, Colloqium Vervoersplanologisch Speurwerk, 1977.

Ruijgrok, C.J., Gedisaggregeerde Keuzemodellen: een evaluatie, Colloqium Vervoersplanologisch Speurwerk, 1978.

Ruijgrok, C.J., and Meijer, H.E.R., De autobeschikbaarheid opnieuw bekeken, IWIS-TNO, januari 1978.

Somermeyer, W.H., Specificatie van Economische Relaties, De Economist, jaargang 115 (1967).

Steenbrink, P.A. Optimization of Transport Networks, John Wiley and Sons Ltd., London, 1974.

Theil, H., Economics and Information Theory, North-Holland Publishing Company, Amsterdam, 1967.

Theil, H., Principles of Econometrics, John Wiley and Sons inc. New York, 1971.

Theil, H., Statistical Decomposition Analysis, North-Holland Publishing Company, Amsterdam, 1972.

Tobin, J., Essays in Economics, volume 2: Consumption and Econometrics, North-Holland/American Elsevier, Amsterdam, 1975.

Warner, S.L., Stochastic Choice of Mode in Urban Travel: a Study in Binary Choice, Evanston: Northwestern University Press, 1962.

Watson, P.L., Homogeneity of Models of Transport Mode Choice: the Dimensions of Trip Length and Journey Purpose, Journal of Regional

Science, Vol. 14, no 2, 1974.

Watson, P.L., Urban Goods Movement, Lexington Books, Toronto/London, 1975.

Watson, P.L., and Westin, R.B., Transferability of Disaggregate Mode Choice Models, Regional Science and Urban Economics 5 (1975), pp. 227-249.

Westin, R.B., Predictions from Binary Choice Models, Journal of Econometrics 2, 1974, pp. 1-16.