SPATIAL DISCRETE CHOICE AND LABOR SUPPLY MODELLING: SOME ALTERNATIVE PROBABILITY CHOICE STRUCTURES

Manfred M. Fischer*
Gunther Maier**

CP-84-51
November 1984

Contribution in the Metropolitan Study: 16

*Department of Geography
University of Vienna
Universitätsstr. 7/V. Stock
A-1090 Vienna - AUSTRIA

**Interdisciplinary Institute for
Urban and Regional Economics
University of Economics
Augasse 2-6
A-1090 Vienna - AUSTRIA

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

CONTRIBUTIONS TO THE METROPOLITAN STUDY:

1. Anas, A. and L.S. Duann (1983) Dynamic Forecasting of Travel Demand. Collaborative Paper, CP-83-45. International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, Austria.

2. Casti, J. (1983) Emergent Novelty and the Modeling of Spatial Processes. Research Report, RR-83-27. IIASA, Laxenburg, Austria.

3. Lesse, P.F. (1983) The Statistical Dynamics of Socio-Economic Systems. Collaborative Paper, CP-83-51. IIASA, Laxenburg, Austria.

4. Haag, G. and W. Weidlich (1983) An Evaluable Theory of a Class of Migration Problems. Collaborative Paper, CP-83-58. IIASA, Laxenburg, Austria.

5. Nijkamp, P. and U. Schubert (1983) Structural Change in Urban Systems. Collaborative Paper, CP-83-57. IIASA, Laxenburg, Austria.

6. Leonardi, G. (1983) Transient and Asymptotic Behavior of a Random-Utility Based Stochastic Search Process in Continous Space and Time. Working Paper, WP-83-108. IIASA, Laxenburg, Austria.

7. Fujita, M. (1984) The Spatial Growth of Tokyo Metropolitan Area. Collaborative Paper, CP-84-03. IIASA, Laxenburg, Austria.

8. Andersson, A.E. and B. Johansson (1984) Knowledge Intensity and Product Cycles in Metropolitan Regions. Working Paper, WP-84-13. IIASA, Laxenburg, Austria.

9. Johansson, B. and P. Nijkamp (1984) Analysis of Episodes in Urban Event Histories. Working Paper, WP-84-75. IIASA, Laxenburg, Austria.

10. Wilson, A.G. (1984) Transport and the Evolution of Urban Spatial Structure. Collaborative Paper, CP-84-41. IIASA, Laxenburg, Austria.

11. Anas, A. (1984) The Combined Equilibrium of Travel Networks and Residential Location Markets. Collaborative Paper, CP-84-42. IIASA, Laxenburg, Austria.

12. Batten, D., P. Newton and J. Roy (1984) Nested Dynamics of Metropolitan Processes and Policies - Melbourne. Collaborative Paper, CP-84-47. IIASA, Laxenburg, Austria.

13. Mackett, R.L. (1984) Nested Dynamics of Metropolitan Processes and Policies - Leeds. Collaborative Paper, CP-84-48. IIASA, Laxenburg, Austria.

14. Dendrinos, D.S. and M. Sonis (1984) Variational Principles and Conservation Conditions in Volterra's Ecology and in Urban Relative Dynamics. Collaborative Paper, CP-84-49. IIASA, Laxenburg, Austria.

15. Batten, D. (1984) The Changing Economic Structure of Metropolitan Regions: A Preliminary Comparative Analysis. Collaborative Paper, CP-84-50. IIASA, Laxenburg, Austria.

16. Fischer, M.M. and G. Maier (1984) Spatial Choice and Labor Supply Modelling: Some Alternative Probability Choice Structures. Collaborative Paper, CP-84-51 IIASA, Laxenburg, Austria.

FOREWORD

Contribution to the Metropolitan Study: 16

The project "Nested Dynamics of Metropolitan Processes and Policies" started as a collaborative study in 1983. This series of contributions is a means of conveying information between the collaborators in the network of the project.

An important analytical starting point for behavioral studies of mobility is random utility theory, which has been a focus of interest in the Regional Issues Project. This paper provides a convenient starting point and state-of-the-art survey for anyone interested in applied mobility analysis and especially analysis of housholds and labor mobility in metropolitan regions.

Ake E. Andersson
Leader
Regional Issues Project
November 1984

1. INTRODUCTION

In the seventies, a wide variety of multiregional economic models has been developed. A comprehensive state of the art review of existing operational models may be found in Issaev et al. (1982). In labour market models both, the demand and the supply side, are taken into account. On the supply side, phenomena such as regional birth and death processes, interregional migration, labour force participation and commuting are modelled. On the demand side, usually the regional stock of capital, the regional production, the demand for goods and services and the regional labour demand are analysed. Most of the labour market models adopt an aggregate modelling approach which usually does not allow detailed conclusions concerning reactions of human actors within the labour market. Furthermore, their behavioural foundation is generally rather poor.

In order to attack these limitations and pitfalls, a micro-oriented perspective has to be pursued. That is, labour markets have to be studied at the scale of individual or disaggregate group preferences. Among several different approaches disaggregate choice theory based on random utility theory is the best developed one. This approach is appealing from both, a theoretical and an empirical point of view. The theoretical appeal comes from the fact that uncertainty in the preferences guiding the choice behaviour is explicitly taken into account and, thus, an appropriate framework for describing inhomogeneities in demand- and supply-side behaviour is provided. The empirical appeal mainly steems from the fact that it enables to link the realm of economic theory, where utility maximization plays a central role with the realm of empirical application. The random utility theory approach has found a wide range of applications (see Bahrenberg, Fischer and Nijkamp, 1984), mostly in transport demand analysis (see Domencich and McFadden, 1975) and most recently also in the area of residential location and housing choice analysis (see e.g. McFadden, 1978, Anas, 1982, Leonardi, 1983).

Changes in regional labour supply may occur as a consequence of labour mobility, like individuals entering or leaving the labour force, changing their residential location or their workplace location, finding a job or becoming unemployed (see Evers and van der Veen, 1983). Adopting the choice-theoretic approach the changes in regional labour supply may be considered as the result of a series of microeconomic decisions among a finite set of discrete alternatives (see Maier and Schubert, 1984). Section 2 attempts to elaborate the theory of utility maximizing discrete choice behaviour and to discuss its implications for modelling the labour supply mobility processes.

Probabilistic discrete choice models are subject to a variety of specification errors which can severly affect the accuracy of model forecasts. Important sources of specification errors refer to

(i) <u>the definition of the choice alternatives</u> (i.e. the problem of defining adequate basic spatial units in the context of spatial choice models),

(ii) <u>the specification of explanatory attributes to be included in the model</u> (missing attributes because of an inadequate theoretical basis or data limitations may seriously bias the estimated coefficients of the included attributes),

(iii) <u>the specification of the form of the utility function</u> (in applied research mainly linear or log-linear functions; the form of the utility function may lead to large differences in the elasticites of the estimated model),

(iv) <u>the assumption underlying the joint distribution of the random terms associated with the choice alternatives</u> (such as the independently and identically distributed multivariate extreme-value distribution, the independently and identically distributed multivariate joint normal distribution or the (generalized) multivariate extreme-value distribution, all leading to different kinds of random utility discrete choice models), and

(v) <u>the specification of the choice structure of the choice problem at hand</u> (such as independent, sequential, recursive sequential and simultaneous structures; in the case of a recursive sequential structure the correct sequence of decisions has to be determined).

Horowitz (1980, 1981) gives some evidence with respect to the second and fourth source of specification errors. Section 3 of this paper pays attention to specification decisions of type (v) within the multinomial logit framework and attempts to determine the effects of using recursive sequential choice structures in a situation where a simultaneous structure is the true one. In order to evaluate if the nested probability choice structures developed in section 2 provide good approximations to an a priori given true simultaneous structure, a Monte Carlo simulation study has been carried out.

2. PROBABILISTIC CHOICE MODELS AND REGIONAL LABOUR SUPPLY

The purpose of this section is to show how the regional labour supply mobility processes may be viewed as a complex decision process consistent with random utility maximizing choice behaviour. In 2.1. the types of labour supply mobility processes are briefly characterized (For a more detailed discussion see Evers and van der Veen, 1983). Section 2.2. reviews the general reasoning behind additive random utility choice models. Particular attention is paid to the multinomial logit model which provides a flexible framework for modelling the labour supply mobility decision process. The discussion in 2.2. will be rather abstract without specific reference to the labour supply context. The reader may, however, find it useful to keep in mind the three labour-supply related mobility decisions to be described in 2.1. Assuming alternative choice structures, in 2.3. several random utility discrete choice models for modelling the regional labour supply mobility processes are developed by adopting the multinomial logit framework.

2.1. Labour supply and discrete decisions

Pursuing the choice-theoretic approach, changes in the regional supply of labour are viewed as the result of a series of decisions among a finite set of mutually exclusive alternatives. If one considers labour supply disaggregated by space, time and categories like qualification a complex system of trajectories and interdependencies emerges. One reason to concentrate on labour supply mobility processes such as changes in employment status, workplace and residential location is the similarity in their speed. They are medium term processes, which typically undergo changes at the individual choice level every several years. Their trajectories are observable over a reasonably long time period.

The evaluation and choice process leading to the <u>residential location decision</u> is mainly based on criteria such as the availability and quality of vacant dwellings, the residential quality of the location and other neighborhood attributes, the accesibility of shops, services and other facilities used on a daily basis, the distance from the workplace, the rent or price of the dwelling etc. (see Leonardi 1983). Such variables may be selected to be included in a spatial choice model of residential location, in which decision makers and choice alternatives (generally geographically defined residential areas) are distributed over space.

The participation or employment status decision is described by the two employment status alternatives participation and non-participation in the labour market. Attributes characterizing these two alternatives reflect the income consequences of the decision, the uncertainty about the chances to actually obtain a job and about a possible future employment status as well as the monetary cost of the decision (see Baumann, Fischer and Schubert, 1984).

The criteria usually postulated to guide the workplace location decision are basically the same as for the participation decision, plus the accesibility of the employment location. As with the residential location decision the choice alternatives are spatially defined.

Obviously, these three decision processes are interlinked. From the point of view of the theory of the rational behaviour they have to be treated as to be taken simultaneously. In empirical applications, however, simultaneity is often unmanageable because the combination of alternatives between the choice sets is likely to make estimation of the choice model very costly and computationally very demanding (see Hensher and Johnson, 1981, 86). Possibilities to overcome this problem will be discussed in section 2.3.

For the purpose of this study it may be assumed that the individuals facing the residential location - employment status - workplace location choice problem are myopic. This means that they consider only the currently available information about the choice alternatives. This also implies that information about the choice alternatives is independent of one's own and other's past choices. A relaxation of this assumption would lead to more realistic but also less tractable choice models, namely to dynamic choice models (see e.g. Fischer and Nijkamp, 1984). The essence of probability discrete choice models and their implications for analysing the regional labour supply mobility process becomes clear without relaxing this simplification.

2.2. Additive random utility choice models: The multinomial logit model

A discrete choice problem is defined by a finite set of disjoint choice alternatives

$$A = \{1, \ldots, A'\} \tag{1}$$

a population of individuals (actors, choice or decision makers)

$$I = \{1, \ldots, I'\} \tag{2}$$

a space Z of observable attributes including measured attributes of alternatives and individual characteristics, a probability density p(z) with $z \varepsilon Z$ characterizing the distribution of attributes in I and a choice probability function $p(a|z)$ specifying the conditional probability of choosing alternative $a \varepsilon A$, given the attributes $z \varepsilon Z$ (see McFadden, 1982, 291). Assuming that there is a priori knowledge of a structure relating z to a, the frequency distribution f(a,z) describing the choice - attribute pairs (a,z) in I can be expressed as

$$f(a,z) = P(a|z)\ p(z) \qquad (a,z) \in (A,Z) \tag{3}$$

where the choice probability function has to satisfy the following conditions

$$o \leqslant P(a\ z) \leqslant 1 \qquad (a,z) \in (A,Z) \tag{4}$$

and

$$\sum_{a \in A} P(a|z) = 1 \qquad z \in Z \tag{5}$$

Spatial discrete choice problems, such as the residential location and the workplace location problem, are characterized by spatially defined alternatives. Note that a decision maker's behaviour may be subject to several external constraints restricting or prohibiting the consideration of certain alternatives. This problem of the nonuniformity of choice sets can be overcome. For the fact that an alternative is not available to one of the individuals can be taken into account, for example, by means of the z-value in such a way that the corresponding choice probability is equal to zero.

Random utility based discrete choice models take for granted that decisions are made at the individual level on the basis of the utility-maximizing principle. It is assumed that there exists a distribution of utility functions from which the utility function U associated with $(i,a)\varepsilon(I\times A)$ is fixed, but latent. In a formal sense, this utility function

$$u_{ia} = U(x_i, y_a, \varepsilon_{ia}) \qquad (i,a)\ \varepsilon\ (I,A) \qquad (6)$$

denotes the (subjective) utility of alternative a to the i-th decision makers where x_i is a vector of attributes characterizing individual i and y_a a vector of attributes describing alternative a with

$$z_{ia} = (x_i, y_a) \qquad (i,a)\ \varepsilon\ (I,A) \qquad (7)$$

The term ε_{ia} represents the random part of utility for alternative a and varies from choice maker to choice maker.

In most random utility discrete choice models utility is assumed to be additively separable (additive utility hypothesis). This assumption allows to formulate (6) as

$$u_{ia} = v_{ia} + \varepsilon_{ia} \qquad (i,a)\ \varepsilon\ (I,A) \qquad (8)$$

where the deterministic component $v_{ia} = V(z_{ia})$ accounts for the effects of the measured attributes of both, the alternatives and the individuals. The random component $\varepsilon_{ia} = \varepsilon(z_{ia},\ldots)$ can represent many types of uncertainty in decision making, such as imperfect information about observation, unobserved constraints that condition individual's choice, unobserved attributes affecting choice and resultant measurement errors as well as other sources of nonhomogeneous or inconsistent behaviour. Since the individual specific variations of ε_{ia} can not be measured, they are treated like random variables which are distributed over I.

A convenient and most widely used functional specification of the systematic component is the linear-in-parameters multiattribute model based on the theory of conjoint measurment (see Timmermans, 1984), leading to

$$u_{ia} = z_{ia}\,\beta + \varepsilon_{ia} \qquad (i,a)\ \varepsilon\ (I,A) \qquad (9)$$

This specification greatly facilitates estimation and statistical interpretation of the parameter vector β which reflects the tastes of the individuals.

The distribution of

$$\varepsilon_{i.} = (\varepsilon_{ia},\ a\ \varepsilon\ A) \qquad (10)$$

conditional to $z_{i.} = (z_{ia} | a\ \varepsilon A)$ is specified as a member of a parametric distributional family

$$F(\varepsilon_{i.} | z_{i.}, \theta) \qquad (11)$$

The utility vector $(u_{ia}, a\ \varepsilon A)$ has a multivariate probability distribution on z_i and β. If this distribution or that of V is known, the choice probability that choice maker i with associated attributes $z_{i.} = (z_{ia'} | a\ \varepsilon A)$ will choose alternative a is then equal to the probability of drawing of a utility vector from this multivariate probability distribution in such a manner that

$$u_{ia} \geq u_{ia'} \qquad a' \neq a,\ a'\ \varepsilon\ A \qquad (12)$$

Thus, the fundamental equation of additive random utility choice models can be formulated as:

$$P(a|z_{i.},\beta) = \text{Prob}((V(z_{ia},\beta) + \mathcal{E}(z_{ia},\beta) \geq V(z_{ia'},\beta) + \mathcal{E}(z_{ia'},\beta);\ a' \neq a,\ a'\ \varepsilon\ A) =: p_{ia} \qquad (13)$$

and for the linear - in - parameters specification as

$$P(a|z_{i.},\beta) = \text{Prob}((z_{ia}-z_{ia'})\beta \geq (\varepsilon_{ia}-\varepsilon_{ia'}); a' \neq a, a' \varepsilon A) \tag{14}$$

The form of the choice probabilities obviously depends on the distributional family F chosen. Different additive random utility choice models result from different distributional families. The <u>mulitnomial probit model</u> (MNP), for example, by assuming that the random terms are independently and identically distributed (IID) according to the multivariate joint normal distribution. While multinomial probit models are theoretically appealing in allowing variation in tastes among choice makers with identical observed attributes as well as the random terms of utility to be correlated and to have unequal variances (see Daganzo, 1979), they become intractable when the number of alternatives is large. This intractability comes from the fact that the functional relationship between the choice probabilites and the measured attributes can not be computed in an analytically closed form. The choice probabilites have to be computed through numerical integration based on various approximation procedures (see e.g. Danganzo et al., 1977).

If F is the <u>IID extreme-value (Weibull or Gnedenko) distribution</u>

$$F(\varepsilon_{i.}|z_{i.},\beta) = \exp(-\sum_{a \varepsilon A} \exp(-\varepsilon_{ia} + \theta_a)) \tag{15}$$

where without loss of generality the parameter θ can be absorbed into the definition of the systematic part of utility (Domencich and McFadden, 1975, 65), the most tractable family of additive random utility discrete choice models, the <u>(linear-in-parameters) multinomial logit (MNL) models</u>, characterized by choice probabilities of the following form

$$P(a|z_{i.},\beta) = \frac{\exp(z_{ia}\,\beta)}{\sum_{a' \varepsilon A} \exp(z_{ia'}\,\beta)} \tag{16}$$

can be derived. A mathematical proof of this derivation can be found in Stopher and Meyburg (1975, 10-11), Hensher and Johnson (1981, 39-42), Anas (1982, 62-63) and Maddala (1983, 60-61).

An appealing feature of the MNL model is that it is computationally much more tractable than the MNP and other random choice models while being a very close approximation to the MNP model since the Weibull distribution is a skewed distribution which closely approximates the normal distribution (see Horowitz, 1980, for an empirical confirmation). Another important characteric is that the relative probabilities of any two alternatives depend only on their systematic components of utility and are independent of other alternatives of the choice set, i.e.

$$\frac{P(a|z_{i.},\beta)}{P(a'|z_{i.},\beta)} = \frac{\exp(z_{ia}\ \beta)}{\exp(z_{ia'}\ \beta)} \tag{17}$$

This property, termed <u>independence from irrelevant alternatives (IIA) property</u>, is both, a strength and a weakness of the MNL model. It is a strength, because it permits the introduction and/or elimination of alternatives in the choice set without reestimating the utility function parameters and, thus, greatly facilitates estimation and forecasting. Furthermore, the IIA property embodies the idea of separability of choices. This separability notion allows to decompose any decision process, such as the residential location - participation - workplace location one, into a series of structurally related decisions (Stopher and Meyburg, 1976, 14-15). It is a limitation when some of the alternatives are similar. The presence of a structure of similarities arising for example from correlation in the unobserved attributes between alternatives invalidates the use of the MNL model. A further feature of the MNL model is worth to be noted, namely that attributes characterizing the decision makers can be included in the utility function but not in a linear-additive manner. These attributes having common values for each alternative would cancel out from the choice probabilites (see Anas, 1982, 64-65).

As a final point in this subsection a few remarks on <u>estimation</u> are in order. The MNL model, i.e. the coefficients of the systematic part of utility, can be estimated via maximum likelihood. Before estimating, certain decisions have to be made with respect to the sampling procedure, which should be used to generate the estimation sample. In the case of <u>random sampling</u> maximum likelihood (ML) estimates can be obtained by one of the standard iterative procedures such as the gradient search, Newton-Raphson, Fisher's scoring, Fletcher-Powell and Davidon procedures. The ML-estimates show all the usual, desirable statistical properties in large samples (see McFadden, 1974). ML-estimation from exogenously stratified samples referring to situations where subsets of individuals with exogenously given attributes are selected, does not yield any new problems as compared to estimation in random sampling. The ML-estimation procedures, however, lead to inconsistent and, thus, asymptotically biased estimates in cases of endogenously stratified (choice-based) sampling. Only recently, progress has been made in developing statistically sound estimators in this case or in the case of more complicated types of sampling processes involving stratification on exogenous and endogenous variables at the same time (see Cosslett, 1981, Manski and McFadden, 1981).

2.3. The multinomial logit framework and some different probability choice structures

The purpose of this section is to develop several regional labour supply mobility discrete choice models within the multinomial logit framework. This will be done by assuming different choice structures underlying the complete residential location - employment status - workplace location choice problem. As mentioned already in 2.2., the separability notion embodied in the IIA-property of the MNL model allows to decompose the regional labour supply mobility process into a series of structurally related decisions and can be used to formulate a variety of decision structures which may be assumed to reflect the true decision behaviour. At least three main choice structures leading to different regional labour supply random utility discrete choice models can be distinguished. These will be described in the sequel.

Suppose that an individual faces three types of decisions: The residential location decision, the participation decision and the workplace decision. Let be

$$A = \{a_1, \ldots, a_{A'}\} \tag{18}$$

the set of spatial options (residential zones) available for the residential location decision,

$$B = \{b_1, \ldots, b_{B'}\} \tag{19}$$

the set of the two employment status alternatives 'participation' ($b = b_1$) and 'non-participation' ($b = b_2$) and

$$C = \{c_1, \ldots, c_{C'}\} \tag{20}$$

the set of spatial options (workplace zones) available for the workplace location decision. For simplicity let us assume that $A' = C'$ and $a_1 = c_1, \ldots, a_{A'} = c_{C'}$. The functional specification of the systematic utility used is the linear-in parameters multiattribute model. Socioeconomic attributes characterizing the decision maker can be included into the utility function, but not in a linear additive way (see Amemiya, 1981, 1491 and Anas, 1982, 64-65). Thus, it is assumed that the systematic component of utility does not depend upon socioeconomic attributes of the decision maker in order to simplify the following discussion.

The first and most simple decision structure is the <u>independent choice structure</u> assuming that there is no dependence of one decision on another. This implies, for example, that the residential location decision has no effect on the participation decision or on the workplace decision. It also implies an additively separable systematic utility of the following form:

$$v_{abc} = y_a \tilde{\beta}_1 + y_b \tilde{\beta}_2 + y_c \tilde{\beta}_3 \qquad a \epsilon A,\ b \epsilon B,\ c \epsilon C \tag{21}$$

where y_a, y_b and y_c represent vectors of attributes characterizing the alternatives in the choice set A, B and C, respectively, and are independent across the decisions. $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ denote the corresponding utility parameter vectors.

Using (16) leads to the following choice probabilities:

$$p_a = \frac{\exp(y_a \tilde{\beta}_1)}{\sum_{a' \epsilon A} \exp(y_{a'} \tilde{\beta}_1)} \tag{22}$$

$$p_b = \frac{\exp(y_b \tilde{\beta}_2)}{\sum_{b' \epsilon B} \exp(y_{b'} \tilde{\beta}_2)} \tag{23}$$

$$P_c = \frac{\exp(y_c \; \hat{\beta}_3)}{\sum_{c' \varepsilon C} \exp(y_{c'} \; \hat{\beta}_3)} \tag{24}$$

The resulting probability choice structure, <u>the (linear-in-parameters IID) independent MNL model</u>, is given by

$$P_{abc} = P_a \; P_b \; P_c \tag{25}$$

Since the attributes describing the three choices are obviously not independent and mutually exclusive (see 2.1.), this model is not acceptable for modelling the labour supply mobility processes.

The second main structure is the <u>recursive sequential or nested choice structure</u>. In this case, it is assumed that the decision process shows a nested (hierarchical) choice structure where results of decisions on lower levels feed into those of higher levels (i.e. feedback effects are taken into account). This implies that the systematic utility is additively separable. One example of an additively separable systematic utility function is:

$$v_{abc} = y_a \; \beta_1 + y_{ab} \; \beta_2 + y_{bc} \; \beta_3 + y_{abc} \; \beta_4 \qquad a \varepsilon A, b \varepsilon B, c \varepsilon C \tag{26}$$

where y_{abc} is the row vector of observed attributes that vary with residential location zone, employment status and workplace location zone (such as journey-to-work travel time), y_{bc} the row vector of observed attributes that vary with employment status and workplace location zone (such as income), y_{ab} the row vector of observed attributes that vary with residential zone and employment status (such as accessibility of kindergartens and day care centers in the residential zone) and y_a the row vector of observed attributes that vary with residential zone only (such as residential quality). β_1, β_2, β_3 and β_4 denote column vectors of utility parameters.

The utility structure (26) is consistent with random utility maximization for the following recursive sequential choice probability structure:

$$P_{abc} = P_a \; P_{b|a} \; P_{c|ab} \tag{27}$$

where $p_{c|ab}$ is the conditional probability of choosing workplace location zone c given the choice of residential location zone a and employment status b, $p_{b|a}$ is the conditional probability of selecting employment status b given the choice of residential zone a and p_a is the marginal probability of choosing residential zone a.

The conditional and marginal probabilities can be seen to be

$$P_{c|ab} = \frac{\exp(y_{bc}\,\beta_3 + y_{abc}\,\beta_4)}{\sum_{c' \in C} \exp(y_{bc'}\,\beta_3 + y_{abc'}\,\beta_4)} \tag{28}$$

$$P_{b|c} = \frac{\exp(y_{ab}\,\beta_2 + (1 - \mu_1)\,M_{ab})}{\sum_{b' \in B} \exp(y_{ab'}\,\beta_2 + (1 - \mu_1)\,M_{ab'})} \tag{29}$$

$$P_a = \frac{\exp(y_a\,\beta_1 + (1 - \mu_2)\,M_a)}{\sum_{a' \in A} \exp(y_{a'}\,\beta_1 + (1 - \mu_2)\,M_{a'})} \tag{30}$$

where the inclusive values M_{ab} and M_a are defined as

$$M_{ab} = \log \sum_{c' \in C} \exp(y_{bc'}\,\beta_3 + y_{abc'}\,\beta_4) \tag{31}$$

$$M_a = \log \sum_{b' \in B} \exp(y_{ab'}\,\beta_2 + (1 - \mu_1)\,M_{ab'}) \tag{32}$$

with μ_1 and μ_2 being coefficients of M_{ab} and M_a, respectively. The inclusive values representing expectations of the outcomes of lower-level choices which might be different among alternatives of higher-level choices serve as linking mechanism in the choice model (27)-(32). From a theoretical point of view, only one elemental alternative is confined to the branch of the decision tree below $b = b_2$.

This three-level nested labour supply mobility discrete choice model belongs to the family of (additive random) nested multinomial logit (NMNL) models, which can be derived from the assumption that the residuals have a generalized extreme-value (GEV) distribution (see McFadden, 1979). As a consequence it is flexible to capture patterns of substitution between alternatives and, thus, relaxes the limiting IIA property of the MNL model. A sufficient and also necessary condition of the three-level nested choice model to be consistent with random utility maximization is that μ_1 and μ_2 lie in the unit interval (McFadden, 1978, 1979).

Estimates of μ_1 and/or μ_2 outside the unit interval can be considered as evidence of a specification error. Finding estimates of μ_1 and μ_2 being equal to one then the feedback effects vanish and (27)-(32) reduces to a model version which is usually termed (three-level) sequential MNL model (c.f. Hensher and Johnson, 1981, 79-81). In this case the choice alternatives within B and C are viewed as distinct alternatives with uncorrelated unobserved attributes. Finding estimates of μ_1 and μ_2 equal to zero the nested model reduces to the simultaneous MNL model version (see Amemiya, 1981). If these parameter estimates are significantly different from one, then there is evidence that the IIA property does not hold (see Maddala, 1983, 77).

The full information maximum likelihood estimation presents some technical problems, but it is possible to use a sequential maximum likelihood procedure and to formulate a sequence of computationally simple maximum likelihood procedures corresponding to the levels of the nested structure. The sequential procedure yields consistent, but not in general efficient estimators (see McFadden, 1981, 242). In the first step β_3 and β_4 are estimated from (28). Then the inclusive values M_{ab} are computed from (31), β_2 and μ_1 are estimated from (29), given the values of M_{ab}. The inclusive values M_a are computed from (32) and β_1 and μ_2 estimated from (30) given the values M_a.

The utility structure (26) is also consistent with utility maximization for the following recursive sequential choice probability structure

$$p_{abc} = p_b \; p_{a|b} \; p_{c|ab} \tag{33}$$

where $p_{c|ab}$ is the conditional probability of choosing workplace location zone c given the choice of residential location zone a and employment status b, $p_{a|b}$ the conditional probability of choosing residential location zone a given the employment status b and p_b the marginal probability of choosing employment status b. The logistic form of these choice probabilities is given by

$$P_{c|ab} = \frac{\exp(y_{bc}\,\beta_3 + a_{abc}\,\beta_4)}{\sum_{c' \in C} \exp(y_{bc'}\,\beta_3 + y_{abc'}\,\beta_4)} \tag{34}$$

$$P_{a|b} = \frac{\exp(y_a\,\beta_1 + y_{ab}\,\beta_2 + (1-\sigma_1)\,S_{ab})}{\sum_{a' \in A} \exp(y_{a'}\,\beta_1 + y_{a'b}\,\beta_2 + (1-\sigma_1)\,S_{a'b})} \tag{35}$$

$$P_b = \frac{\exp((1-\sigma_2)\,S_b)}{\sum_{b' \in B} \exp((1-\sigma_2)\,S_{b'})} \tag{36}$$

with the inclusive values

$$S_{ab} = \log \sum_{c' \in C} \exp(y_{bc'}\,\beta_3 + y_{abc'}\,\beta_4) \tag{37}$$

$$S_b = \log \sum_{a' \in A} \exp(y_{a'}\,\beta_1 + y_{a'b}\,\beta_3 + (1-\sigma_1)\,S_{a'b}) \tag{38}$$

From a theoretical point of view, only A' elemental alternatives are confined to the branch of the decision tree below $b = b_2$.

The two discrete choice models (27)-(32) and (33)-(38) differ in that the first one assumes a three-level nested multinomial logit model with the employment status decision conditional on the residential location decision, whereas the second one assumes a three-level nested logit model with the residential location choice conditional on the employment status decision. In the first choice model version the inclusive value parameter μ_1 measures the correlation among the random terms due to workplace zones similarity within a residential zone and for the same employment status. An estimate of μ_1 equal to zero denotes no correlation among unobserved workplace location related attributes whereas an estimate of μ_1 equal to one indicates nearly identical unobserved attributes for the workplace zones. Similarly, the inclusive value parameter μ_2 is a measure of correlation among unobserved employment status related attributes within a residential zone with an estimate of μ_2 equal to one indicating nearly identical unobserved random attributes for the employment status alternatives and an estimate of μ_2 equal to zero denoting no correlation among the employment status alternatives. In the second three-level nested choice model the parameters σ_1 and σ_2 have a similar interpretation. σ_1 measures workplace zone similarity in unobserved random attributes whereas σ_2 measures residential zone similarity in unobserved random attributes (see Anas, 1982, 137-138).

Now let us turn to the third structure, the simultaneous choice structure. In this case no assumption is made about the independency or the nested structure of the choice process. Rather it is assumed that an individual makes all three choices simultaneously. The simultaneous probabilistic choice structure corresponding to this case with utility structure (26) is given by

$$P_{abc} = \frac{\exp\,(y_a\beta_1 + y_{ab}\beta_2 + y_{bc}\beta_3 + y_{abc}\beta_4)}{\sum\limits_{a'\epsilon A}\ \sum\limits_{b'\epsilon B}\ \sum\limits_{c'\epsilon C} \exp(y_{a'}\beta_1 + y_{a'b'}\beta_2 + y_{b'c'}\beta_3 + y_{a'b'c'}\beta_4)} \tag{39}$$

From a theoretical point of view the set of alternatives is confined to A'C' + A' different alternatives. That is, an elemental alternative is either a combination of residential zone a, employment status $b = b_1$ (e.g. 'participation') and workplace zone c or a combination of residential zone a, employment status $b = b_2$ (e.g. 'non-participation') and workplace zone c with workplace zone c and residential zone a coinciding.

Obviously the simultaneous choice structure leads to a complex discrete choice model in which there are likely to be serious estimation problems in the case of a large number of alternatives. This probability choice structure, however, leads to the best utility parameter estimates, if the underlying choice structure is a simultaneous one. In the context of labour supply mobility decisions it seems to be the most appropriate one. Whereas from a behavioural point of view the independent choice structure is totally implausible, a recursive sequential decision structure may be viewed as acceptable if the IIA assumption of the joint MNL model is significantly invalid and/or if simultaneity becomes unmanagable (i.e. large A' and C'). If the recursive sequential choice structuring approach is adopted it is obviously necessary to determine the most adequate nested structure. For the estimated parameters seem to be sensitive to the sequence of decisions chosen.

Finally, it is worthwile to note that the simultaneous and the three-level recursive sequential choice structures can be combined in such a way that two-level nested regional labour supply mobility discrete choice models result. Suppose a two-level nested choice structure where - as in (28)-(33) - first the residential location is chosen and then simultaneously the employment status and the workplace location, given the residential location. Choice alternatives at the lower level of the choice structure are combinations of workplace zones c and the employment status 'participation' plus the alternative 'non-participation'. Assuming utility structure (26) the following NMNL model is obtained:

$$P_{abc} = P_a \; P_{bc|a} \tag{40}$$

$$P_{bc|a} = \frac{\exp(y_{ab}\,\beta_2 + y_{bc}\,\beta_3 + y_{abc}\,\beta_4)}{\sum_{b'\in B}\sum_{c'\in C}\exp(y_{ab'}\,\beta_2 + y_{b'c'}\,\beta_3 + y_{ab'c'}\,\beta_4)} \tag{41}$$

$$P_a = \frac{\exp(y_a\,\beta_1 + (1-\omega)\,W_a)}{\sum_{a'\in A}\exp(y_{a'}\,\beta_1 + (1-\omega)\,W_{a'})} \tag{42}$$

where the inclusive value W_a is defined as

$$W_a = \log\Big(\sum_{b'\in B}\sum_{c'\in C}\exp(y_{ab'}\,\beta_2 + y_{b'c'}\,\beta_3 + y_{ab'c'}\,\beta_4)\Big) \tag{43}$$

where the inclusive value parameter ω measures the workplace zone - employment status similarity in unobserved random attributes.

The second two-level NMNL model differs from (40)-(43) in that it assumes a choice structure with the simultaneous choice of residential and workplace location conditional on the employment status decision. The sequential probability choice structure consistent with maximizing (26) is given by

$$P_{abc} = P_b \; P_{ac|b} \tag{44}$$

with

$$P_{ac\,b} = \frac{\exp(y_a\,\beta_1 + y_{ab}\,\beta_2 + y_{bc}\,\beta_3 + y_{abc}\,\beta_4)}{\sum_{a'\in A}\sum_{c'\in C}\exp(y_{a'}\,\beta_1 + y_{a'b}\,\beta_2 + y_{bc'}\,\beta_3 + y_{a'bc'}\,\beta_4)} \tag{45}$$

A'C' alternatives (combinations of residential and workplace zones) are confined to the branch below $b = b_1$, whereas only A' spatial options (i.e. workplace zone and residential zone coinciding) are confined to the branch below $b = b_2$. The utilities of the remaining (A'C' - A') alternatives confined to $b = b_2$ are a priorily set equal to minus infinity. Consequently, the corresponding conditional choice probabilities (45) become zero.

$$P_b = \frac{\exp((1-\tau)\ T_b)}{\sum_{b' \varepsilon B} \exp((1-\tau)\ T_{b'})} \tag{46}$$

where the inclusive value T_b is defined as

$$T_b = \log\left(\sum_{a' \varepsilon A} \sum_{c' \varepsilon C} \exp(y_{a'}\ \beta_1 + y_{a'b}\ \beta_2 + y_{bc'}\ \beta_3 + y_{a'bc'}\ \beta_4\right) \tag{47}$$

The inclusive value parameter τ measures the residential and workplace location similarity in unobserved random attributes.

3. <u>THE ACCURACY OF RECURSIVE SEQUENTIAL PROBABILITY CHOICE STRUCTURES AS AN APPROXIMATION TO A SIMULTANEOUS CHOICE STRUCTURE</u>

While the simultaneous probability choice structure is appealing and often the most adequate one of the alternative choice structures, it becomes unmanageable in the case of a large number of spatial options. Thus it seems to be interesting and useful to determine the effects of using a recursive nested probability choice structure in situations where a simultaneous structure may be valid. It is the main purpose of this section to present some evidence by means of a Monte Carlo simulation study and to analyse, whether

- o the three-level nested multinomial logit regional labour supply mobility discrete choice models (26)-(31), shortly termed <u>NMNL model 1</u> and (32)-(37), termed <u>NMNL model 2</u>, and
- o the two-level nested multinomial logit regional labour supply mobility discrete choice models (40)-(43), <u>NMNL model 3</u>, and (44)-(47), <u>NMNL model 4</u>,

provide good approximations to an a priori given true simultaneous probability choice structure. Furthermore, it is investigated if the simultaneous multinomial logit model (39) is able to reproduce this choice structure.

For simplicity it is assumed that only three spatial options are available for both, the residential and the workplace location choice. The utility function is specified in accordance with (26) as

$$u_{abc} = y_a^1 \; {}^o\beta_1^1 + y_{ab}^1 \; {}^o\beta_2^1 + y_{bc}^1 \; {}^o\beta_3^1 + y_{abc}^1 \; {}^o\beta_4^1 + \varepsilon_{abc} \tag{48}$$

with

y_a^1 an observed attribute (such as residential quality), which varies with the residential location zone only,

y_{ab}^1 an observed attribute (such as disutility of working), which varies with both, the residential zone and the employment status,

y_{bc}^1 an observed attribute (such as income measured as wage-income in the case of b_1 and as non-wage-income in the case of b_2), that varies with both, the employment status and the workplace zone

y_{abc}^1 an observed attribute (such as journey-to-work travel time) that varies with the residential location zone, the employment status and the workplace location zone.

${}^o\beta_1^1, {}^o\beta_2^1, {}^o\beta_3^1, {}^o\beta_4^1$ utility parameters, whose values are a priori given

ε_{abc} IID Weibull distributed random terms.

In order to evaluate the accurracy of the nested probability choice structures as approximations to a simultaneous choice structure the measurements of y_a^1, y_{ab}^1, y_{bc}^1 and y_{abc}^1 are defined as follows:

$$y_a^1 = \quad y_1^1 = 2.0,\ y_2^1 = 1.5,\ y_3^1 = 1.0$$

$$y_{ab}^1 = \begin{cases} y_{11}^1 = 0.6,\ y_{21}^1 = 0.4,\ y_{31}^1 = 0.5 \\ y_{12}^1 = 0.1,\ y_{22}^1 = 0.0,\ y_{32}^1 = 0.1 \end{cases} \tag{49}$$

$$y_{bc}^1 = \begin{cases} y_{11}^1 = 2.0,\ y_{21}^1 = 0.5,\ y_{12}^1 = 1.5 \\ y_{22}^1 = 0.5,\ y_{13}^1 = 3.0,\ y_{23}^1 = 0.5 \end{cases}$$

and

$$y^1_{abc} = \begin{cases} y^1_{111} = 0.0,\ y^1_{211} = 2.0,\ y^1_{311} = 3.0 \\ y^1_{121} = 0.0,\ y^1_{221} = 0.0,\ y^1_{321} = 0.0 \\ y^1_{112} = 2.0,\ y^1_{212} = 0.0,\ y^1_{312} = 1.5 \\ y^1_{122} = 0.0,\ y^1_{222} = 0.0,\ y^1_{322} = 0.0 \\ y^1_{113} = 3.0,\ y^1_{213} = 1.5,\ y^1_{313} = 0.0 \\ y^1_{123} = 0.0,\ y^1_{223} = 0.0,\ y^1_{323} = 0.0 \end{cases} \tag{49 cont.}$$

The true probability choice structure is defined by a fixed, but arbitrary parameter vector

$$\Phi^0 = ({}^0\beta^1_1 = 1.5,\ {}^0\beta^1_2 = -2.5,\ {}^0\beta^1_3 = 1.2,\ {}^0\beta^1_4 = -0.5) \tag{50}$$

For convenience, let D denote the set of 12 elemental alternatives available at the lowest level of the decision trees (see fig.1). Using (39) the theoretical choice probabilities $p_d(\Phi^0)$ with $d \in D$ are calculated. For a sample of I' = 10.000 individuals random utility values are generated for each of the 12 alternatives. The ε-terms are obtained by means of a standard random number generator producing identically and independently uniform zero-one distributed pseudo random numbers $\tilde{\varepsilon}$ which are transformed into IID Weibull distributed pseudo random numbers ε as follows (see Zielinski,1978)

$$\varepsilon = -\log(-\log(\tilde{\varepsilon})) \tag{51}$$

Next, for each individual of the sample the alternative associated with the highest value of random utility is identified in order to calculate observed frequencies of choice of the sample. The evaluation of the model's performance is based on both, the concordance of the estimated utility parameters with the corresponding true parameter values as well as the fit of predicted and theoretical frequencies of choice. In order to prevent the results of the analysis from being influenced by random fluctuations in frequencies of choice being not relevant for the question considered, R=30 replications are made.

Estimates of the utility parameters of the simultaneous MNL model are obtained by full information maximum likelihood estimation, whereas the sequential maximum likelihood estimation approach is taken in the case of the nested choice model versions. The maximization procedure utilized is a combination of a Newton-Raphson variable metric routine and a gradient search technique adopted from Wegener and Graef (1982). This procedure utilizes the different characteristics of the two techniques. The gradient search technique generally has a high probability of convergence, but tends to be slow on flat parts of the likelihood function. The Newton-Raphson technique usually is much faster, but may diverge in the case of bad starting points.

Presenting the results, let us first focus on the question, whether the simultaneous MNL model (39) is able to reproduce the a priori given choice structure. The range and the average of the resulting ML parameter estimates are summarized in table 1. The average values fit the true parameter values almost exactly. With the exception of the β_1^1's the magnitudes of the 30 parameter estimates are quite stable. The numbers in parentheses denote the standard errors of the ML-estimates.

In order to test, if

$$\hat{\Phi} = (\hat{\beta}_1^1, \hat{\beta}_2^1, \hat{\beta}_3^1, \hat{\beta}_4^1) \tag{52}$$

is statistically different from Φ^0 Hotelling's T^2 test, a generalization of the conventional t-test, may be used. The T^2 statistic is defined as

$$T^2 = (\hat{\Phi} - \Phi^O) \, \Sigma_{\hat{\Phi}}^{-1} \, (\hat{\Phi} - \Phi^O)' \tag{53}$$

where $\Sigma_{\hat{\Phi}}$ denotes the estimated variance-covariance matrix of $\hat{\Phi}$. This statistic is distributed with an F value:

$$F = T^2 \, (I' - Q)/(I' - 1) \, Q \tag{54}$$

with Q and I'-Q degrees of freedom, where I' is the number of individuals in the sample and Q the number of parameters in the model (see Anas 1982, 152). The T^2-statistic values indicate that the parameters estimates of the simultaneous MNL model are not significantly different from the true parameter values (see table 1).

Another test to evaluate the accuracy of the simultaneous MNL model concerns the model's ability to predict the actual choices. For this purpose there is a variety of goodness of fit measures available, such as the sum of squared residuals, the sum of squared residuals weighted by estimated probabilites or adjusted by the true probabilities, the simple correlation coefficient between observed and predicted data or measures based on the log-likelihood function. For a discussion of these and other measures see Amemiya (1981, 1503-1506), Domencich and McFadden (1975, 122-125).

In this study the 'mean absolute percentage error' statistic defined as

$$K^O = \frac{\sum_{r=1}^{R} \sum_{d \in D} |I' \; \pi_d^r(\hat{\Phi}) - I' \; p_d(\Phi^O)|}{\sum_{r=1}^{R} \sum_{d \in D} I' \; p_d(\Phi^O)} \; 100 \tag{55}$$

is used, where $\Pi_d^r(\hat{\Phi})$ denotes the relative frequency of alternative d obtained from the ML-estimated simultaneous MNL model in the r-th replication and $p_d(\Phi^0)$ the true choice probability of d. This statistic measures the average deviation of the predicted from the theoretical frequencies of choice.

The resulting K^O-value of the model is 1.5606 meaning that the predicted frequencies in the average deviate only about one-and-a-half percent from the true (theoretical) frequency. Thus, the true frequency is almost perfectly reproduced by the simultaneous model. Consequently, both the T^2-test and the K^o-statistic lead to the conclusion that the simultaneous MNL model performs extremly well to the a priori defined simultaneous choice structure situation.

Now let us turn to the question whether the four nested multinomial logit regional labour supply mobility discrete choice models provide good approximations to the a priori defined true probability choice structure. The ML-estimates of the utility and inclusive value parameters obtained by the sequential ML-procedure are presented in terms of range, average value and standard error (see table 2). The β-coefficients all have the expected sign. The β_3^1- and β_4^1-ML-estimates come rather close to the true values, whereas the ranges of the β_1^1- and β_2^1-ML-estimates are much larger. Nevertheless, their average values fit quite well, almost perfectly in the case of model versions 2 and 4. This indicates that the results are sensitive to the sequence of the decisions even if only to a rather low degree.

Furthermore, the T^2-test can be applied to determine whether the maximum likelihood vectors of the NMNL models differ significantly from the true parameter values. In this case T^2 is given by

$$T^2 = (\hat{\phi}^n - \phi^{on}) \; \Sigma_{\hat{\phi}^n}^{-1} \; (\hat{\phi}^n - \phi^{on})' \tag{56}$$

where $\hat{\phi}^n$ denotes the ML-vector of the nested choice model version n (n = 1,...,4) and ϕ^{on} the corresponding true parameter vector. For n = 1,2 the ML-vector contains two, for n = 3,4 one inclusive value parameter estimate in addition to the β-estimates. ϕ^{on} is defined as follows

$$\phi^{on} = \begin{cases} (1.5,\ -2.5,\ 1.2,\ -0.5,\ 0.0,\ 0.0) & \text{for } n=1,2 \\ (1.5,\ -2.5,\ 1.2,\ -0.5,\ 0.0) & \text{for } n=3,4 \end{cases} \tag{57}$$

The results outlined in table 3 clearly demonstrate that the T^2 test does not reject the models as approximations to the true simultaneous choice structure. The inclusive parameter estimates do not significantly differ from zero.

The mean absolute percentage error statistic K^o as a measure of the prediction performance of the nested choice models can be defined in a way analogous to (55), with the relative frequencies of choice of the ML-estimated simultaneous model replaced by the corresponding relative frequencies of choice obtained from the ML-estimated NMNL models, i.e. $\Pi_d^r(\hat{\Phi})$ by $\Pi_d^r(\hat{\Phi}^n)$. The statistics summarized in table 3 demonstrate the high prediction performance of the models. From both aspects, the statistical reliability of the parameters and the prediction performance, all four nested models can be considered as generally acceptable approximations to the true simultaneous probability choice structure. The two-level nested choice models are slightly preferable to the three-level nested ones. Overall considered model 4 appears to be the most appropriate approximation.

The K^o-statistic can be modified in order to compare the predicted choices of the NMNL models with those obtained from the simultaneous MNL model. The modification is as follows: The frequencies of the simultaneous model, i.e. $I'\Pi_d^r(\hat{\Phi})$, are replaced by those of the nested version, i.e. $I'\Pi_d^r(\hat{\Phi}^n)$, while the theoretical frequencies, i.e. $I'p_d(\Phi^0)$, are substituted by the frequencies computed from the simultaneous model, i.e. $I'\Pi_d^r(\hat{\Phi})$. This statistic may be termed K^1. The results, outlined in table 3, indicate that the nested models are good approximations to the simultaneous one.

It is worthwhile to mention that the independent as well as sequential regional labour supply mobility discrete choice models can by no means be considered as acceptable simplifications of the true probability choice structure in cases where a simultaneous choice structure is valid. This is not very surprising from a methodological point of view.

4. SUMMARY AND CONCLUSIONS

This paper attempts to elaborate the theory of additive utility maximizing random utility discrete choice behaviour in the context of regional labour supply mobility. From a choice theoretic point of view regional labour supply is considered as the result of a series of decisions. Specific emphasis is laid upon three types of regional labour supply mobility, namely on employment status decisions (non-spatial discrete choice), on workplace and on residential location decisions (spatial discrete choices). Evidently, these three choices are interrelated. According to the neoclassical theory of rational behaviour they have to be treated as to be simultaneously taken.

Since in empirical applications simultaneity often becomes unmanageable due to large numbers of alternatives, it is both interesting and useful to determine the effects of choosing alternative probability choice structures (such as recursive sequential structures) in a situation, where a simultaneous one may be valid. This analysis is carried out in the multinomial logit framework and evidence is presented with respect to three-level and two-level nested linear-in-parameter multinomial logit models. In order to evaluate whether these nested probability choice structures provide good approximations to an a priori given simultaneous structure 30 replications of the following approach are carried out: For a sample of 10.000 individuals IID Weibull distributed random numbers are generated. Random utilities and true choice probabilities for each of 12 elemental alternatives are calculated using a priori given utility parameter values and observations. Given this sample (with 30 replications) the utility and the inclusive value parameters of the nested logit model versions are estimated by the sequential ML-estimation approach. The results obtained are presented in terms of range, average and standard error of the ML - estimates. Hotelling's T^2-test is applied to determine, if the parameter estimates differ significantly from their true values. The mean absolute percentage error statistic is used to measure the model's prediction performance. Both tests clearly indicate that the nested choice model versions appear to be generally acceptable approximations to the true simultaneous structure, whereas neither the independent nor the sequential multinomial logit model versions can be considered as acceptable simplifications.

A clear implication of these results is to use nested multinomial logit models instead of independent or sequential ones in situations where a simultaneous structure might be valid, but is computationally unmanageable. Then the effects of simplification generally can be kept rather small. There is also some weak evidence that two-level nested models are preferable to three-level nested approximations in the context of regional labour supply mobility.

REFERENCES

Amemiya, T., 'Qualitative response models: A survey', Journal of Economic Literature, vol.19 (1981), pp. 1483-1536.

Anas, A., Residential Location Markets and Urban Transportation. Economic Theory, Econometrics, and Policy Analysis with Discrete Choice Models, Academic Press, New York et al., 1982.

Bahrenberg, G., Fischer, M. M. and Nijkamp, P. (eds.) Recent Developments in Spatial Data Analysis: Methodology, Measurement, Models, Gower, Aldershot, 1984 (in press).

Baumann, J. H., Fischer, M. M. and Schubert, U., 'Multiregional labour market modelling and regionalization: Some further results', Paper presented at the 2nd World Conference of Regional Science, Rotterdam, 1984.

Cosslett, S. R., 'Efficient estimation of discrete - choice models', in: Manski, C. F. and McFadden, D. (eds.), Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge (Mass.) and London, 1981, pp. 51-111.

Daganzo, C., Multinomial Probit. The Theory and its Application to Demand Forecasting, Academic Press, New York et al., 1979.

Daganzo, C., Bouthelier, F. and Sheffi, Y., 'Multinomial probit and qualitative choice: A computationally efficient algorithm', Transportation Science, vol.11 (1977), pp. 338-358.

Domencich, T. A., and McFadden, D., Urban Travel Demand: A Behavioral Analysis, North-Holland, Amsterdam and American Elsevir, New York, 1975.

Evers, G. H. M. and van der Veen, A., 'A theoretical framework for analysing labor supply-mobility: An application of spatial discrete choice modelling', Paper presented at the 23rd European Congress of the Regional Science Association, Poitiers, 1983.

Fischer, M. M. and Nijkamp, P., 'Explanatory discrete spatial data and choice analysis: A state-of-the-art review', Research Memorandum, Department of Economics, Free University, Amsterdam, 1984.

Hensher, D. A. and Johnson, L. W., Applied Discrete-Choice Modelling, Croom Helm, London and Wiley, New York, 1981.

Horowitz, J., 'The accuracy of the multinomial logit model as an approximation to the multinomial probit model of travel demand, Transportation Research, vol. 14 B (1980), pp. 331-341.

Horowitz, J., 'Identification and diagnosis of specification errors in the multinomial logit model', Transportation Research, vol. 15 B (1981), pp. 345-360.

Issaev, B., Snickars, F., Nijkamp, P. and Rietveld, P., Multiregional Economic Modeling: Practice and Prospect, North-Holland, Amsterdam, New York and Oxford, 1982.

Leonardi, G., 'A choice-theoretical framework for household mobility and extensions', in: Final Report of the Research Project on Feasibility Analysis of Studies on Transport-Location Interactions, Rome, National Council, 1983.

Maddala, G. S., Limited - Dependent and Qualitative Variables in Econometrics, Cambridge University Press, Cambridge et al., 1983.

Maier, G. and Schubert, U., 'Struktur und Probleme ökonometrischer interregionaler Arbeitsmarktmodelle - mit einigen ausgewählten Ergebnissen für Österreich', in: Bahrenberg, G. and Fischer, M. M. (eds.) Theorie und quantitative Methodik in der Geographie, Bremen, University of Bremen press, 1984 (in press).

Manski, C. F. and McFadden, D., 'Alternative estimators and sample designs for discrete choice analysis', in: Manski, C. F. and McFadden, D. (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge (Mass.) and London, 1981, pp. 2-50.

McFadden, D., 'Conditional logit analysis of qualitative choice behaviour', in: Zarembka, P. (ed.), Frontiers in Econometrics, Academic Press, New York and London, 1974, pp. 105-142.

McFadden, D., 'Modelling the choice of residential location', in: Karlqvist, A., Lundqvist, L., Snickars, F. and Weibull, J. W. (eds.), Spatial Interaction Theory and Planning Models, North Holland, Amsterdam, 1978, pp. 75-96.

McFadden, D., 'Qualitative methods for analysing travel behaviour of individuals: Some recent developments', in: Hensher, D. A. and Stopher, P. R. (eds.), Behavioural Travel Modelling, Croom Helm, London, 1979, pp. 279-318.

McFadden, D., 'Econometric models of probabilistic choice', in: Manski, C. F. and McFadden, D. (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge (Mass.) and London, 1981, pp. 198-272.

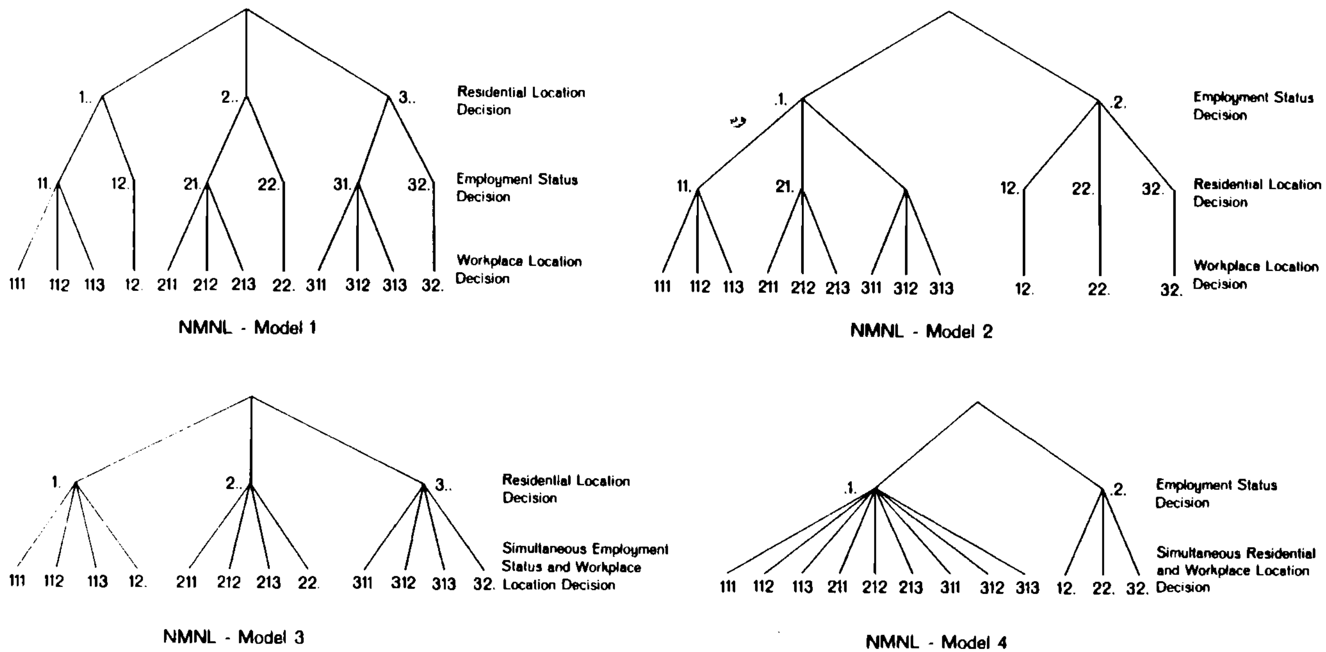
Stopher, P. R. and Meyburg, A. H., 'Behavioural travel-demand models', in: Stopher, P. R. and Meyburg, A. H., Behavioural Travel-Demand Models, Lexington Books, Lexington and Toronto, 1976, pp. 3-53.

Timmermans, H. J. P., 'Decision models for predicting preferences among multiattribute choice alternatives', in Bahrenberg, G., Fischer, M. M. and Nijkamp, P. (eds.), Recent Developments in Spatial Data Analysis: Methodology, Measurement, Models, Gower, Aldershot, 1984, pp. 337-354.

Wegener, M. and Graef, F., 'Estimating multiattribute spatial choice models', WP-82-93, International Institute for Applied Systems Analysis, Laxenburg, 1982.

Zielinski, R., Erzeugung von Zufallszahlen: Programmierung und Test auf Digitalrechnern, VEB Fachbuchverlag, Leipzig, 1978.

FIGURES AND TABLES



Note: The first digit refers to the residential location, the second to the employment status and the third to the workplace location decision.

Fig. 1: Alternative probability choice structures for labour supply modelling

Table 1: Parameter estimates of the simultaneous multinomial logit model: Range, average, standard error[a], reliability[b] and prediction performance[c] (sample size: I' = 10000, R = 30)

Parameter	Estimate[a]			Reliability[b]
	max	mean	min	
β_1^1	1.571	1.508 (0.033)	1.434	0.055
β_2^1	-2.079	-2.514 (0.020)	-2.658	0.017
β_3^1	1.237	1.204 (0.020)	1.140	0.032
β_4^1	-0.465	-0.503 (0.012)	-0.524	0.079
Prediction performance				
K^o	1.5606			
Log-likelihood	21903.			

[a] standard errors in parentheses [b] measured in terms of the T^2 statistic
[c] measured in terms of the K^o statistic

Table 2: The effects of nested probability choice structures upon parameter estimates: Range, average and standard error[a] (sample size I' = 10000, R = 30)

	3-Level Nested Versions						2-Level Nested Versions					
Parameter	NMNL-Model 1			NMNL-Model 2			NMNL-Model 3			NMNL-Model 4		
	max	mean	min	max	mean	min	max	mean	min	max	mean	min
β_1^1	1.846	1.476 (0.160)	1.232	1.649	1.511 (0.074)	1.335	1.969	1.614 (0.180)	1.381	1.582	1.510 (0.036)	1.440
β_2^1	-1.929	-2.620 (0.354)	-3.218	-2.132	-2.501 (0.268)	-3.033	-2.054	-2.506 (0.116)	-2.729	-2.127	-2.535 (0.144)	-2.922
β_3^1	1.242	1.202 (0.020)	1.138	1.242	1.202 (0.020)	1.138	1.246	1.203 (0.021)	1.137	1.245	1.202 (0.021)	1.137
β_4^1	-0.466	-0.503 (0.012)	-0.527	-0.466	-0.503 (0.012)	-0.027	-0.467	-0.504 (0.012)	-0.525	-0.467	-0.504 (0.012)	-0.525
μ_1	0.147	0.035 (0.057)	-0.066		—			—			—	
μ_2	0.383	0.024 (0.235)	-0.629		—			—			—	
σ_1		—		0.213	0.006 (0.100)	-0.192		—			—	
σ_2		—		1.230	0.325 (0.337)	-0.149		—			—	
ω		—			—		0.176	-0.153 (0.210)	-0.562		—	
τ		—			—			—		0.093	-0.018 (0.059)	-0.204

[a] standard errors in parentheses

Table 3: The effects of nested probability choice structures upon parameter estimates: Reliability[a] and prediction performance[b] (sample size I' = 10000, R = 30)

	3-Level Nested Versions		2-Level Nested Versions	
	NMNL-Model 1	NMNL-Model 2	NMNL-Model 3	NMNL-Model 4
parameter	reliability[a] of the parameter estimates			
β_1^1	0.023	0.021	0.396	0.081
β_2^1	0.114	0.00002	0.002	0.059
β_3^1	0.010	0.010	0.015	0.012
β_4^1	0.077	0.077	0.093	0.089
μ_1	0.379	———	———	———
μ_2	0.010	———	———	———
σ_1	———	0.004	———	———
σ_2	———	0.931	———	———
ω	———	———	0.531	———
τ	———	———	———	0.089
prediction performance[b]				
K^0	1.8890	1.8301	1.7019	1.6919
K^1	0.9916	0.9230	0.5844	0.5382
Log likelihood	21902	21902	21903	21903

[a] measured in terms of the T^2 statistic

[b] measured in terms of the K^0 and K^1 statistics