Recent developments regarding similarities in transport modelling

Nadine Schuessler

Kay W. Axhausen

STRC 2007   **August 2007**

STRC 2007

Recent developments regarding similarities in transport modelling

Nadine Schuessler
IVT
ETH Zürich
CH-8093 Zürich
phone: +41-44-633 30 85
fax: +41-44-633 10 57
schuessler@ivt.baug.ethz.ch

Kay W. Axhausen
IVT
ETH Zürich
CH-8093 Zürich
phone: +41-44-633 39 43
fax: +41-44-633 10 57
axhausen@ivt.baug.ethz.ch

August 2007

Abstract

Overcoming the independence of irrelevant alternatives (IIA) property of the basic multinomial logit (MNL) model is a major research issue in the field of discrete choice modelling. In recent years, several approaches have been developed to achieve this goal with different degrees of appropriateness.

On the one hand there are very flexible models, which are able to account for complex correlation structures and a wide variety of interdependencies between alternatives by opening the variance-covariance structure. But they require a lot of effort in terms of specification and computation. On the other hand there are less complex models, which introduce a similarity factor into the systematic part of the utility function, which decreases the utility of an alternative with respect to its similarity with other alternatives. They are easier to estimate and applicable to large choice sets. However, these models were designed to solve specific, in particular route choice, problems and not offhand transferable.

This paper summarises and evaluates different approaches to overcome the IIA property. Special consideration is given to approaches that are easy to compute and applicable to a combined route, mode and destination choice model.

Keywords

Discrete choice, IIA property, similarities

1   Introduction

Today, discrete choice models have manifold applications. They are used in a wide variety of contexts to simulate consumer choice. They base on the idea that a decision-maker is confronted with a set of discrete alternatives and has to choose one of them. The model itself estimates for each alternative the probability of being chosen assuming that a decision-maker seek to maximise his or her utility. The utility depends on the decision-maker's individual preferences, the choice situation, the characteristics of the alternative and its similarities with the other available alternatives. The underlying utility function is split into two elements: a systematic part and a random part, for both of which the analyst has to make suitable distributional assumptions. In principle, the analyst is free from a priori constraints in his or her choice of approach to capture similarities among alternatives. Similarities can be included in the systematic part through a suitably measure of similarity or through an appropriate specification of the variance-covariance matrix of the random part.

Either approach, with different degrees of appropriateness, overcomes the "Independence of Irrelevant Alternatives" (IIA) property of the basic Multinomial Logit (MNL) model. This property implies that the ratio of the choice probabilities of any two alternatives is not affected by the availability or the attributes of other alternatives and is therefore independent of the size and structure of the choice set. Not correcting for the IIA property leads in many cases to - very - misleading model results and forecasts.

Recent research has developed several approaches to overcome this structural problem, but none is completely satisfactory. On the one hand there are very flexible models, which are able to describe complex correlation structures. They account for a wide variety of interdependencies between alternatives by opening the variance-covariance structure of the model. But they require a lot of effort in terms of specification and computation and are not obviously suitable for large sets of overlapping alternatives. This is especially a problem in the field of transport research. Realistic problems addressed here are often characterised by large sets of alternatives as for example in route or destination choice. For models that simultaneously address two or more choices this problem increases even further, think for example of a joint destination and route choice model.

On the other hand there are less complex models, which are easier to estimate and applicable to large choice sets. However, these were designed to solve specific problems and are not offhand transferable. They introduce a similarity factor to the systematic part of the utility function, which decreases the utility of an alternative with respect to its similarity with other alternatives. The majority of these models has been developed for route choice, though some applications for destination and location choice have been proposed. Nevertheless, there have hardly been any approaches for dealing with multi-modal situations or other combined choice problems.

This paper summarises and evaluates different approaches to overcome the IIA property. First, an introduction to the MNL model and its IIA property and to the idea of the different approaches is given. Subsequently, the approaches that subdivide alternatives into nests are introduced before those changing the variance-covariance structure and the ones that introduce similarity factors in the deterministic part of the utility function are presented. The paper concludes with a discussion of the different approaches and an outlook for future research. Special consideration is given to question if they are applicable to a combined route, mode and destination choice model.

2 The MNL model and its IIA property

Discrete choice models are a standard for modelling consumer behaviour. In transport research, they are used for all aspects of travel behaviour, including, but not limited to household activity scheduling, destination choice, route choice and mode choice. Therefore, discrete choice models are of special importance for the evaluation of transport policies, such as infrastructure investment or setting of tolls.

In a discrete choice model, an individual - the decision-maker - is confronted with a set of discrete alternatives, the choice set, from which he or she has choose one. As a decision rule, it is assumed that the decision-maker seeks to maximise his or her personal utility. The utility of each alternative is characterised by its measurable attributes captured by the deterministic component V_{in} of the utility function. Beyond that, there are utility components that cannot be measured directly due to several reasons. First, there is heterogeneity of preferences across decision-makers. Second, the knowledge and the information processing abilities of decision-makers are limited. Third, there are further uncertainties regarding the choice process, including attributes which the analyst is not able or not resourced to measure. These elements are usually represented by the random term ε_{in} of the utility function. Thus, the following utility function is postulated:

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

with V_{in} being defined as $V_{in} = f(\beta, x_{in})$, where β is a vector of taste coefficients, and x_{in} a vector of the attributes of alternative i as faced by respondent n in the specific choice situation. In addition, socio-demographic attributes of respondent n can be included in the systematic part of the utility function.

The discrete choice model itself estimates for each alternative the probability of being chosen from a given choice set:

$$P(i|C_n) = P[U_{in} \geq U_{jn}, \forall j \in C_n] \tag{2}$$

The most commonly used discrete choice model is the Multinomial Logit Model (MNL) (McFadden, 1974). It is based on the assumption that the random terms, often called error terms, are identically and independently (i.i.d.) Gumbel distributed. The choice probability of each alternative i can then be calculated as:

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_j e^{\mu V_{jn}}} \tag{3}$$

Thereby, μ is related to the standard deviation of the Gumbel variable ($\mu^2 = \frac{\pi}{6\sigma^2}$), where, in

the absence of a heterogeneous population, μ is generally constrained to a value of 1.

The advantages of the MNL model are its flexibility in terms of its deterrence sensitivity, and the ease of the parameter estimation (c.f. Ben-Akiva and Lerman (1985)). On the other hand, the MNL model has several disadvantages, the most prominent being the property of the "Independence from Irrelevant Alternatives" (IIA property). This property states: The relative ratio of the choice probabilities of two alternatives does not depend on the existence or the characteristics of other choice alternatives.

$$\frac{P(i|C_n)}{P(k|C_n)} = \frac{\frac{e^{\mu V_{in}}}{\sum_j e^{\mu V_{jn}}}}{\frac{e^{\mu V_{kn}}}{\sum_j e^{\mu V_{jn}}}} = e^{\mu(V_{in} - V_{kn})} \tag{4}$$

An illustration for this problem is the well-known red bus/blue bus paradox (Debreu, 1960), which describes two mode choice situations. First, the decision-maker is facing two alternatives: taking the car or a red bus. It is assumed, that each alternative has a choice probability of 50%. In the second scenario, a blue bus with the same attributes relevant for the decision as the red bus is added to the choice set. Because the new alternative is just another option for using public transport, one would expect, that the share of the additional alternative comes completely at the expense of the red bus and the resulting choice probabilities should be: $P_{Car} = 50\%$, $P_{redbus} = 25\%$ and $P_{bluebus} = 25\%$ – ignoring for now the potential mode shift because of increased frequencies on the bus network. However, because of the IIA property, the MNL returns the same choice probability for each alternative ($P_{Car} = 33\%$, $P_{redbus} = 33\%$ and $P_{bluebus} = 33\%$) to guarantee that the ratio between the probabilities for the car and the red bus stays equal to one.

Though the red bus and the blue bus obviously share a lot of characteristics and are therefore similar, the MNL model ignores this completely. The same applies for any other choice context, as demonstrated for example by Daganzo and Sheffi (1977) for private transport route choice, where the similarity between routes is derived from their overlap. Thus, one possible interpretation of the IIA property of the MNL model is its failure account for similarities between alternatives. Mathematically, similarities can be represented by correlations. Since the error terms in the MNL model are independently distributed, no correlations are included in the model as can be seen from the variance-covariance matrix for 5 alternatives depicted below, where because of its immanent symmetry, only the upper triangle of the matrix is shown. The matrix consists only of the variances of the alternatives' utilities. The covariances are all equal to zero.

$$
\begin{pmatrix}
\sigma_{11}^{in} & 0 & 0 & 0 & 0 \\
 & \sigma_{22}^{in} & 0 & 0 & 0 \\
 & & \sigma_{33}^{in} & 0 & 0 \\
 & & & \sigma_{44}^{in} & 0 \\
 & & & & \sigma_{55}^{in}
\end{pmatrix}
$$

This property can lead to biased parameter estimates. Furthermore, the model misses an important aspect of the actual choice behaviour. Solving this issue is still an ongoing research topic as is the question whether similarities between alternatives have positive or negative effects on their choice probabilities. Theory postulates that similarities reduce the probability to be chosen. However, recent studies such as Hoogendoorn-Lanser and Bovy (2007), Frejinger and Bierlaire (2007) or van Eggermond *et al.* (2007) suggest that this assumption does not hold for all choice contexts.

3   Accounting for similarities between discrete choice alternatives

Before exploring the different approaches of how to handle similarities in discrete choice modelling, one has to reflect, what kind of similarities can appear in real choice situations, as they can differ enormously depending on the decision context. Considering for example mode choice, similarities between private or public transport modes can lie in characteristics such as accessibility, comfort or levels of privacy. Correlations between private transport routes appear, if routes share links, whereas similarities among public transport connections are determined through comparable time slots or journey times, equal interchange facilities or the same operator. An individual facing a destination choice problem has to deal with even more possible similarities. Those can include the geographic region, the landscape, the journey direction (route overlap), the weather, the products/services offered or shared parking sites - to name only a few.

There are different ways of describing similarities between alternatives. Three general approaches can be distinguished that will be described in the following sections:

- subdividing alternatives into nests,

- opening the variance-covariance structure, and

- introducing similarity factors in the deterministic part of the utility function.

The first group basically contains the Generalised Extreme Value (GEV) models other than the MNL model. The alternatives are subdivided into groups, the so-called nests. Correlations may remain within the nests, but between the nests they are eliminated. In the classic Nested Logit (NL) model the nests are completely disjoint whereas in the Cross Nested Logit (CNL) model each alternative can belong to more than one nest. However, though particularly the CNL is able to represent nearly all kinds of correlations, a realistic nesting structure is highly complex and therefore cumbersome to estimate.

Most of recent research efforts have however focussed on the second group of models, more specifically on the Mixed Multinomial Logit (MMNL) Models. Inspired by the Multinomial Probit Model where multivariate normally distributed error terms replace the i.i.d. Gumbel distributed ones of the MNL model, in an MMNL model the deterministic part of the utility function is re-formulated while the i.i.d. Gumbel error terms remain. A multivariate randomly distributed error term is introduced that captures similarities which cannot be modelled deterministically. This leads to models that are able to account for any kind of correlation structure and taste heterogeneity. However, these models require a lot of effort in terms of specification

and computation and are thus still hardly applicable to choice situations with large numbers of alternatives.

The models of the third group aim to capture correlation effects by correcting the systematic component of the utility function. Based on the implicit availability/perception model (IAP) presented by Cascetta *et al.* (1996) they rest upon the assumption that the utility of an alternative is decreased by its degree of similarity with other alternatives. Thus, they add a deterministic similarity measure to the utility function. Again, the deterministic part is reformulated and the error terms remain i.i.d. Gumbel distributed. The crucial aspect of these approaches is the appropriate choice of the similarity factor. Alternatives in transport choice problems are usually characterised by attributes of different variable type. Hence, a suitable similarity measures has to cope with different variable types, such as dichotomous, qualitative and quantitative variables.

Gower (1985) provides the definition of a general measure of similarity as well as a description of its properties and its applicability to different types of variables. In his measure the similarity between two alternatives is determined by comparing each of their attributes, assigning a score for the degree of similarity and combining those scores to a single coefficient. Below, a short description of similarity scores for single-level variables of different variable types is followed by the introduction of Gower's general similarity measure that combines them.

Dichotomous and qualitative variables can be treated in a similar way. While dichotomous variables represent the presence or absence of a characteristic, qualitative variables embody qualities of equal standing. The similarity score for those variables can only be unity or zero depending on the question if they are equal or not. Care must however be taken with regard to the question if the absence of both attributes of a dichotomous variable is considered as equality or not. The similarity or dissimilarity between quantitative variables can be measured through the distance between their characteristics. This distance is then converted into a fraction between zero and unity to get the score for the similarity coefficient.

Gower (1985) then combines the similarity scores $s_k(x_{ik}, x_{jk})$ of the individual variables to the "General Coefficient of Similarity" between the alternatives i and j, which has the following functional form

$$S_{ij} = \frac{\sum_{k=1}^{p} w_k(x_{ik}, x_{jk}) \cdot s_k(x_{ik}, x_{jk})}{\sum_{k=1}^{p} w_k(x_{ik}, x_{jk})} \tag{5}$$

Thereby the function imposed for $s_k(x_{ik}, x_{jk})$ can differ for each variable. $w_k(x_{ik}, x_{jk})$ represents the weight put on the individual attribute to reflect the importance of a variable or its reliability.

With the help of this approach even hierarchies of alternatives can be compared in the way that

the similarity score $S_{ij}^{(k)}$ of the secondary alternatives associated to attribute k is imposed as weight for the primary variable (c.f. Gower (1985), p. 400):

$$S_{ij} = \sum_{k=1}^{p} S_{ij}^{k}(x_{ik}, x_{jk}) \cdot s_k(x_{ik}, x_{jk}) \tag{6}$$

Any depth of nesting can be represented by generalising this procedure.

4   Generalised Extreme Value models

The family of Generalised Extreme Value (GEV) models was first introduced by McFadden (1978) for the context of residential location choice with the aim to allow for dependencies between the unobserved error-terms of the alternatives while maintaining the closed-form of the MNL model. McFadden (1978) demonstrated that the MNL model is also a GEV model. The most popular model, apart from the MNL, of this family is the Nested Logit model, first presented by Ben-Akiva (1973, 1974). Since the Nested Logit model is not able to capture all kinds of correlations, the Cross Nested Logit model was introduced by McFadden (1978) and refined by Ben-Akiva and Bierlaire (1999). It was further generalised to the Generalised Nested Logit model (Wen and Koppelman, 2001) and the Network GEV model (Bierlaire, 2002). All these models are briefly presented in the following.

4.1   The Nested Logit model

The basic idea of the Nested Logit model is to divide all alternatives of a choice set into disjoint groups, the so-called nests. Correlations may remain within the nests, but between the nests they are eliminated. Thus, the entire utility function for alternative i belonging to nest C_{nm} has to be reformulated. The systematic component is split into two parts and incorporates the alternative specific effects V_{in}' as well as the impacts associated with the nest $V_{C_{mn}}$:

$$U_{in} = V_{in}' + \varepsilon_{in} + V_{C_{mn}} + \varepsilon_{C_{mn}} \tag{7}$$

ε_{in} and $\varepsilon_{C_{mn}}$ are independent. The distribution of the error-term ε_{in} remains IID Gumbel with a scale parameter σ_k, while the error-terms $\varepsilon_{C_{mn}}$ jointly follow a generalised extreme-value distribution in a way that the random variable $max_{j \in C_{mn}} U_{jn}$ is Gumbel distributed with scale parameter μ. Each nest C_{mn} has a composite utility $V_{C_{mn}}'$, also called *expected maximum utility* or *LOGSUM*:

$$V_{C_{mn}}' = V_{C_{mn}} + \frac{1}{\mu_m} ln \sum_{j \in C_{mn}} e^{\mu_m V_{jn}} \tag{8}$$

where $V_{C_{mn}}$ is the utility common to all alternatives in nest C_{mn}. Thus, the probability of choosing alternative i that is part of nest C_{mn} from the individual choice set C_n can be calculated as the product of the probability, that nest C_{mn} is chosen from the set of all nests and the probability that alternative i is chosen from the alternatives belonging to nest m:

$$P(i|C_n) = P(C_{mn}|C_n) \cdot P(i|C_{mn}) \tag{9}$$

with

$$P(C_{mn}|C_n) = \frac{e^{\mu V'_{C_{mn}}}}{\sum_{l=1}^{M} e^{\mu V'_{C_{ln}}}} \tag{10}$$

and

$$P(i|C_{mn}) = \frac{e^{\mu_m V_{in}}}{\sum_{j \in C_{mn}} e^{\mu_m V_{jn}}} \tag{11}$$

For $\frac{\mu}{\mu_m} = 1\ \forall k$ the NL collapses to the MNL model.

As such, correlation between the error-terms of alternatives nested together is introduced. However, the model does not capture potential correlations between nests. This can be illustrated by the variance-covariance matrix. The example here shows the covariances for 5 alternatives, of which alternative 1 and 2 belong to the same nest, alternative 3 to a second one and the alternatives 4 and 5 to the third one:

$$\begin{pmatrix} \sigma_{11}^{in} & \sigma_{12}^{in} & 0 & 0 & 0 \\ & \sigma_{22}^{in} & 0 & 0 & 0 \\ & & \sigma_{33}^{in} & 0 & 0 \\ & & & \sigma_{44}^{in} & \sigma_{45}^{in} \\ & & & & \sigma_{55}^{in} \end{pmatrix}$$

4.2   The Cross Nested Logit model

A solution to the problem of missed correlations between alternatives that do not belong to the same nest is the Cross-Nested Logit (CNL) model. Recently, it became very popular. Among those who applied it in a transport context are Small (1987), who's Ordered Generalised Extreme Value model is mathematical identical to the CNL, Ramming (2002) for departure time choice, Vovsha and Bekhor (1998) and for route choice and Bierlaire *et al.* (2001) for mode choice. In the CNL model, each alternative can belong to more than one nest. To represent the degree of membership to a nest, an allocation parameter $0 \leq \alpha_{im} \leq 1$ is introduced. Thus, the utility function of the CNL (Ben-Akiva and Bierlaire, 1999) can be formulated as:

$$U_{in} = V'_{in} + \varepsilon_{in} + V_{C_{nm}} + \varepsilon_{C_{nm}} + ln \quad \alpha_{im} \tag{12}$$

with ε_{in} and $\varepsilon_{C_{nm}}$ being defined the same way as for equation 7. The choice probability for an alternative i has then to be calculated over all nests m it partially belongs to:

$$P(i|C_n) = \sum_{m=1}^{M} P(C_{mn}|C_n) \cdot P(i|C_{mn}) \tag{13}$$

It is important to note, that any functional relationship can be defined for the allocation parameter α_{im} depending on the choice context, though often, simple point-estimates are used. Abbe *et al.* (2007) derived a normalisation for α_{im} which is:

$$\sum_m \alpha_{im}^{\frac{\mu}{\mu_m}} = c, \forall j \in C_n \tag{14}$$

where c is a constant that does not depend on i.

Thus, the CNL is theoretically able to depict all kinds of correlation structure by allowing the error-terms of alternatives that are somehow nested together to be correlated. The following example shows the matrix for a five alternative, three nests example, with the following membership structure: alternative 1 belongs to nest a, alternative 2 to nests a and b, alternative 3 to nest b, alternative 4 to nests b and c and alternative 5 to nest c.

$$\begin{pmatrix} \sigma_{11}^{in} & \sigma_{12}^{in} & 0 & 0 & 0 \\ & \sigma_{22}^{in} & \sigma_{23}^{in} & \sigma_{24}^{in} & 0 \\ & & \sigma_{33}^{in} & \sigma_{34}^{in} & 0 \\ & & & \sigma_{44}^{in} & \sigma_{45}^{in} \\ & & & & \sigma_{55}^{in} \end{pmatrix}$$

However, in practice the CNL model soon leads to highly complex structures, which make it difficult to specify and computationally hard to estimate.

Since Ben-Akiva and Bierlaire (1999), the CNL model has been extensively scrutinised. Papola (2004) for example conjectured that the covariance of the utilities of two different alternatives depends not only on the inclusion into one nest, but on the degree of membership of both alternatives to the same nest and implied a linear relation between the NL correlations and the CNL correlations. He validated this assumption starting with the observation that the CNL model is a generalisation of a two-level NL model. However, Abbe *et al.* (2007) demonstrated that this supposition is only a fairly good approximation for the limited case, when the CNL model has only bipolar shared alternatives.

4.3 The Generalised Nested Logit model

To derive a more general formulation of the CNL, Wen and Koppelman (2001) proposed the Generalised Nested Logit (GNL) model. It summarises the NL, CNL and other NL derivates such as the Paired Combinatorial Logit model by Koppelman and Wen (2000) through normalisation of the CNL model structure. The GNL model fractionally assigns each alternative to an nest and different LOGSUMs can be calculated for each nest. The normalisation of Wen and Koppelman (2001) was formally proved by Abbe *et al.* (2007). Furthermore, the latter give

guidance to derive a CNL model from any arbitrary variance-covariance structure.

4.4 The Network GEV model

Another approach of generalisation was proposed by Bierlaire (2002) with the Network GEV (NGEV) model. The author showed that any correlation structure represented by a network with certain properties can be modelled with a NGEV model and furthermore, that every such model is indeed a GEV model. Thereby, the properties for the network are straightforward: The network is not allowed to include circuits, it has to have one root node without predecessors, the alternatives have to be represented by leafs without successors, and each node in the network has to be part of a continuous path between the root and one alternative. This model formulation is especially appealing because of its intuitive way of capturing even complex correlation structures. It eases the formulation of a complex model by its recursive definition.

5   Probit and Mixed Multinomial Logit models

Recently, it has been popular to develop discrete choice models that overcome the IIA property by opening the variance-covariance structure. The most extensive of these models is the Multinomial Probit model. Mixed Multinomial Logit models aim to combine Probit-like error terms with the closed-form of the MNL model. Both model forms are briefly described in the following. Subsequently several studies that successfully employed MMNL models to transport choice problems are presented.

5.1   The Probit model

In the Probit model, discussed for example by (Daganzo, 1979), multivariate Normal distributed error terms replace the i.i.d. Gumbel distributed ones of the MNL resulting in the most general variance-covariance structure:

$$\begin{pmatrix} \sigma_{11}^{in} & \sigma_{12}^{in} & \sigma_{13}^{in} & \sigma_{14}^{in} & \sigma_{15}^{in} \\ & \sigma_{22}^{in} & \sigma_{23}^{in} & \sigma_{24}^{in} & \sigma_{25}^{in} \\ & & \sigma_{33}^{in} & \sigma_{34}^{in} & \sigma_{35}^{in} \\ & & & \sigma_{44}^{in} & \sigma_{45}^{in} \\ & & & & \sigma_{55}^{in} \end{pmatrix}$$

Thus, any variance-covariance structure can be specified and all kinds of correlation structures between the alternatives of the choice set can be depicted. Probit model have for example been applied by Yai *et al.* (1997) or Daganzo and Sheffi (1977) to private transport route choice problems. The respective authors specified models, in which the covariances of the route utilities are proportional to the length of link overlaps. However, the formulation of the probit model is complex and its utility function does not have a closed form. It can not be solved analytically and requires simulation for estimation as well as application. Thus, it is only applicable if the number of parameters and alternatives is small.

5.2   The Mixed Multinomial Logit models

The family of Mixed Multinomial Logit (MMNL) or Logit Kernel (LK) models was proposed by Walker (2001) with the aim to combine the advantages of a Probit model with those of a Logit model. Two conceptually different but mathematically identical approaches of the MMNL model exist: the Error-Components Logit (ECL) and the Random-Coefficients Logit (RCL). In the ECL model the systematic part V_{in} of the utility function is split into a deterministic component V_{in}' and a random error component η_{in}. η_{in} is generally assumed to be multivariate Normal distributed. Correlation is introduced by allowing some alternatives to

share the same error component. The RCL model accommodates unobserved variation across individuals in their sensitivity to observed exogenous variables by specifying some entries of the vector β in the equation $V_{in} = f(\beta, x_{in})$ to be random variables. This can also be represented by adding a multivariate Normal distributed variable η_{in} to the utility function. Thus, for the ECL as well as the RCL model the utility function can be denoted as follows:

$$U_{in} = V'_{in} + \eta_{in} + \varepsilon_{in} \tag{15}$$

V'_{in} is re-formulated using only the attributes of the alternatives, the decision maker and the choice situation while η_{in} is a multivariate randomly distributed error term that captures similarities which cannot be modelled deterministically.

In the present context, the ECL approach is of special interest. By specifying the structure of the error-components such that a given set of alternative shares an error-component, correlation between these alternatives is allowed for. In theory, the resulting model can approximate any correlation structure, including heteroscedastic ones, arbitrarily closely. As such, the model can also replicate the variance-covariance matrix of the general Probit model. Like the Probit model, the ECL model has the disadvantage that simulation is required in estimation and application. In addition, imposing the right identification restrictions so that a unique solution can be obtained from the infinite set of optimal solutions of the unconstrained model is a difficult and time-consuming business and an often overlooked one as argued in Walker (2002). If the MMNL model additionally allows for random taste variation (e.g. in an RCL framework), these problems go much further because before identification issues can be solved, the appropriate distribution function for the random parameters has to be determined. This altogether makes the model difficult to be applied in large-scale forecasting systems. For further discussion see Walker (2002), Walker *et al.* (forthcoming) and Ben-Akiva and Bolduc (1996).

Yet, several studies have successfully applied MMNL models to transport related choice problems. Private transport route choice problems have for example been examined by Bekhor *et al.* (2001), Ramming (2002) and Frejinger and Bierlaire (2007). Guo and Bhat (2005) presented an MMNL model for residential location choice while Hess *et al.* (2005) modelled mode choice employing an MMNL model.

Bekhor *et al.* (2001) and Ramming (2002) used the 1997 Transportation Survey of Faculty and Staff conducted by the MIT Planning Office to estimate MMNL models for 188 observations with a maximum of 51 and a median of 30 route alternatives. They assume that the covariances are proportional to free flow travel time of path overlaps.

Frejinger and Bierlaire (2007) on the other hand developed a different approach. Correlation is not established using link overlaps but so-called subnetwork components. A subnetwork component is a subsection of the route network consisting of a continuous sequence of links

that are easily identifiable and behaviourially relevant. Subnetwork components can either be derived from the network hierarchy or from route descriptions in personal interviews. Paths sharing a network component are assumed to be correlated even if they are not physically overlapping. Thus, correlation is rather defined from a behavioural point of view. The authors tested different model specifications with subnetworks based on a data set containing 2978 observations from for cars in the city of Borlaenge, Sweden. The choice set size ranges from 2 to 43 alternative path with a majority of choice sets containing less that 15 paths.

In addition, it is interesting to note that Ramming (2002) as well as Frejinger and Bierlaire (2007) include a Path Size factor (c.f. section 6.2) to explicitly account for further correlations between alternatives. In both cases the MMNL models with Path Size factor outperform those without. Yet, all MMNL models result in better model fits than the basic MNL model with Path Size factor.

The Mixed Spatially Correlated Logit (MSCL) model suggested by Guo and Bhat (2005) combines an MMNL model with a Paired Generalised Nested Logit model. It has been developed for residential location choice. The PGNL structure accounts for correlations between adjacent spatial units whereas the mixing Normal distribution captures unobserved taste heterogeneity. The approach was applied to model the residential location choice of 236 households within parts of Dallas County for zones of different sizes and characteristics. The authors found that the Mixed GEV combining a closed-form correlation structure with an open-form account for taste variations resulted in a good model-fit and at the same time computational efficiency compared to a pure MMNL model.

A Mixed GEV model was also applied by **?** to model data from a Stated Preference long-distance mode choice survey in Switzerland. The aim of the survey was to estimate the hypothetical demand for a new transport system in Switzerland, the so-called Swiss Metro (c.f. Abay (1999) (in German) and Bierlaire *et al.* (2001)). Nested Logit and Cross Nested Logit models (c.f. section 4.1 and 4.2) are combined with Normal distributed random terms to capture taste heterogeneity. The results emphasised, that there is a significant risk of confounding effects of taste heterogeneity and correlation since these two phenomena are not necessarily clearly distinguishable. This is especially pointed out by the difficulties the authors experienced with the estimation of the Mixed CNL model that were only partly due to the model complexity.

6   Including similarity factors in the utility function

The inclusion of a similarity factor in the systematic part of the utility function can be derived from the implicit availability/perception model (IAP) presented in Cascetta *et al.* (1996). They state that an individual is not able to consider all alternatives of the universal choice set because of the individual's imperfect knowledge and his or her limited information processing abilities. Whether an alternative is evaluated depends directly on its availability and the decision-maker's perception. The alternatives considered are a subset of the universal choice set: the individual choice set C_n of decision-maker n. For decision-maker n the degree of membership of an alternative i to the choice set C_n is represented by the variable ϑ_{in}. The probability of choosing an alternative from the universal choice set hence depends on the probability that this alternative is included in the individual choice set, and the choice probability of the alternative within this restricted choice set. With the re-formulated utility function

$$U_{in} = V'_{in} + \vartheta_{in} + \varepsilon_{in} \tag{16}$$

the new choice probability for alternative i is

$$P(i|C_n) = \frac{e^{\mu(V'_{in} + ln\vartheta_{in})}}{\sum_j e^{\mu(V'_{jn} + ln\vartheta_{jn})}} \tag{17}$$

Again, V'_{in} is reformulated and ε_{in} remains i.i.d. Gumbel distributed. The crucial aspect of these approaches is the appropriate choice of ϑ_{in}. Since there are still no correlations between the error terms, the variance-covariance matrix is the same as for the MNL model. All similarities are captured as systematic attributes of the alternatives.

All the approaches below are based on this general idea but differ in the specification of ϑ_{in}. In the majority of cases, ϑ_{in} represents the independence of alternative i from all other alternatives of the universal choice set. Thereby, the dependency of two alternatives is usually equivalent to their degree of equality. Except for the Competing Destinations model by Fotheringham (1988), the underlying assumption is that the independence of an alternative increases its probability to be perceived as a separate alternative. However, recent empirical evidence shows that this assumption can not be hold in every case. There are situations where similarities between alternatives are perceived positive and this effect outweighs the statistical effect.

6.1   C-Logit

When establishing the IAP model, Cascetta *et al.* (1996) proposed also a way to account for similarities in private transport route choice, the so-called C-Logit model. The Commonality Factor CF_{in} indicates the percentage of route length that route i shares with other routes by

comparing the total length of route i with the length of the overlapping links. Cascetta *et al.* (1996) proposed four different formulations for CF_{in}:

$$CF_{in} = -\beta_{CF} \cdot ln \sum_{j \in C_n} \left(\frac{L_{ij}}{\sqrt{L_i \cdot L_j}} \right)^{\gamma} \tag{18}$$

$$CF_{in} = -\beta_{CF} \cdot ln \sum_{a \in \Gamma_i} \frac{l_a}{L_i} N_{an} \tag{19}$$

$$CF_{in} = -\beta_{CF} \cdot \sum_{a \in \Gamma_i} \frac{l_a}{L_i} ln N_{an} \tag{20}$$

$$CF_{in} = -\beta_{CF} \cdot ln \left[1 + \sum_{j \in C_n, i \neq j} \left(\frac{L_{ij}}{\sqrt{L_i \cdot L_j}} \right) \left(\frac{L_i - L_{ij}}{L_j - L_{ij}} \right) \right] \tag{21}$$

where β and γ are coefficients, that have to be estimated, L_{ij} is the length of links shared by i and j, Γ_i the set of links of route i, l_a the length of link a, and N_{an} number of links using link a. Thus, the choice probability of route i is

$$P(i|C_n) = \frac{e^{\mu(V'_{in} + CF_{in})}}{\sum_j e^{\mu(V'_{jn} + CF_{in})}} \tag{22}$$

Unfortunately, there is no theoretical guidance to which of these formulations should be used. Cascetta *et al.* (1996) themselves applied Equation 19 whereas Equation 18 is used by Ramming (2002) and Cascetta *et al.* (2002), and Equation 21 is evaluated by Prato and Bekhor (2007) and Ramming (2002).

Vrtic (2003) used the formulation in Equation 18 and combined it with a Nested Logit model to the Nested C-Logit (NCL) model. The NCL model was developed for a simultaneous route and model choice model. The nesting structure accounts for similarities between private transport and public transport alternatives respectively. Deterministic correlations within the nests on the other hand are captured by the Commonality Factor that depicts a route's similarity with other routes in terms of the route length.

6.2 Path Size Logit

The Path Size (PS) Logit model of Ben-Akiva and Bierlaire (1999) was also developed for private transport route choice problems. The length of each route is corrected by the so-called

path size PS_{in}. Only a distinct route, i.e. a route with no overlaps with other routes, can get the maximum path size of one. Path sizes different from one are calculated based on the length of the links within the route i and the length of the routes that share a link with it relative to the length of the shortest route using the link. Accordingly, the choice probability for route i is

$$P(i|C_n) = \frac{e^{\mu(V'_{in}+lnPS_{in})}}{\sum_j e^{\mu(V'_{jn}+lnPS_{in})}} \tag{23}$$

Ben-Akiva and Bierlaire (1999) propose two different formulation for PS_{in}, the first one being

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \delta_{aj}} \tag{24}$$

where Γ_i is the set of all links of path i, l_a is the length of link a, and L_i the length of path i. δ_{aj} equals one if link a is on path i and zero otherwise. The second formulation additionally accounts for the relative ratio between the length of the shortest path $L^*_{C_n}$ in C_n and the length of all paths using link a.

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \delta_{aj} \frac{L^*_{C_n}}{L_j}} \tag{25}$$

Ramming (2002) states that this model formulation has a major shortcoming: Its second term is not affected by the length of other then the shortest route if a link is used by more than one route. Thus, he derived a General Path Size (GPS) factor. He reformulates the second part of Ben-Akiva and Bierlaire's Path Size factor to account for the contribution of the individual links. The basic idea is to give each link the size one and to allocate this size among the routes using that link. The size of a route is then calculated as the sum of its link sizes weighted according to the length of the route compared with other routes sharing that link. The influence of this weighting is given by the size allocation parameter γ.

$$GPS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} (\frac{L_i}{L_j})^\gamma \delta_{aj}} \tag{26}$$

Especially for large γ, Ramming (2002) achieved the best model results for $\gamma = \infty$, this formulation assigns the size of a shared ling primarily to the shortest path using that link.

However, Hoogendoorn-Lanser *et al.* (2005), who applied the PS factor and the GPS factor to multi-modal route choice, as well as Frejinger and Bierlaire (2007) found the interpretation of this approach difficult. In contrast to the original PS factor that can be interpreted as an approximation of the variance-covariance matrix, the GPS factor introduces asymmetry into the model by explicitly favouring the shortest route. In addition, the empirical analysis of the GPS factor showed that it captures part of the explanatory power of the variables related to the

units the GPS factor is measured in. Furthermore, Hoogendoorn-Lanser *et al.* (2005) expressed the need to have a close look at the value of γ before applying it and to explicitly estimate β_{PS}, which had been fixed to one by Ramming (2002) and Ben-Akiva and Bierlaire (1999).

6.3 Competing Destinations

Another derivation of the IAP model is provided by Fotheringham (1988) with the Competing Destinations (CP) model. He states that the probability for an alternative to be chosen by the decision-maker depends on its probability to be included in his or her choice set. This probability depends on its similarity with other alternatives and can be specified in different ways according to the decision-context. Mathematically, the CP approach is identical with the IAP model with the choice probability for alternative i being

$$P(i|C_n) = \frac{e^{\mu(V'_{in}+lnCP_{in})}}{\sum_j e^{\mu(V'_{jn}+lnCP_{jn})}} \tag{27}$$

However, the CP model bases on the assumption that close geographic proximity to other stores increases the probability of a store to be included in the decision-maker's choice set instead of decreasing it. Fotheringham (1988) suggests two specifications for the context of consumer store choice. The first formulation, taken from Fotheringham (1983), sums up the distances d_{ij} of a store i to all other stores j in the universal choice set C containing I stores, weighted by the utility of the other stores j.

$$CP_{in} = \left(\frac{1}{I-1}\sum_{j,j\neq i}\frac{V'_{in}}{d_{ij}}\right)^{\theta} \tag{28}$$

The second formulation, originally been proposed by Borgers and Timmermans (1987), simply accounts for the average distance of store i to all other stores.

$$CP_{in} = \left(\frac{1}{I-1}\sum_{j,j\neq i}d_{ij}\right)^{\theta} \tag{29}$$

Both measures result in lower values for spatially isolated stores, thus, the probability of an alternative to be chosen is decreased.

6.4 Prospective utilities

Kitamura (1984) also prosecuted the aim to refine destination choice modelling. He stated that dependencies between destinations not only exist in spatial but also in temporal and causal

dimensions. His argumentation is in line with the recent activity based research that emphasises the importance of trip chains for destination choice. Thus, Kitamura (1984) developed a destination choice model that accounts for trip chaining effects by introducing a factor of Prospective Utility (PU) U_{jn} into the systematic part of the utility function:

$$U_{in} = V'_{in} + PU_{in} + \varepsilon_{in} = V'_{in} + \sum_j q_{jn}(U_{jn} \quad \theta \quad d_{ij}) + \varepsilon_{in} \tag{30}$$

with q_{jn} being the subjective probability that decision-maker n carries out an activity in zone j after his activity in zone i, d_{jn} the spatial distance between i, and j and θ the disutility parameter for d_{ij}. PU_{in} can be interpreted as a measure of perceived accessibility of zone i. It can be modified to account for different trip purposes and due to it recursiveness also for longer trip chains.

6.5 The Independence of a Connection

Most similarity factors developed so far have been either designed for private transport route choice or for spatial choices. Only little attention has been paid to public transport connection or multi-modal route choice. Exceptions are Hoogendoorn-Lanser and Bovy (2007), Cascetta and Papola (2003) and Friedrich *et al.* (2001). Whereas most similarity factors focus on a spatial dimension of similarity, the influence of the spatial dimension is less decisive for public transport connection choice and mainly restricted to shared transfer points. Instead, temporal aspects are highly relevant, especially for inter-urban public transport. While Hoogendoorn-Lanser and Bovy (2007) found the use of the same public transport leg to be the significant description for overlaps in multi-modal route choice, Cascetta and Papola (2003) demonstrated that correlations between departure times are much stronger than those between the same public transport modes. Another important aspect for public transport similarity measure is the fare.

Thus, Friedrich *et al.* (2001) designed a similarity measure specifically for public transport connection choice, the Independence of a Connection (IND) factor. It enters the systematic part of the utility function and thus the choice probability of alternative i.

$$P(i|C_n) = \frac{e^{\mu(V'_{in} + ln \quad IND_{in})}}{\sum_j e^{\mu(V'_{jn} + ln \quad IND_{jn})}} \tag{31}$$

IND is defined as the reciprocal of the sum of similarities of alternative i with all other alternatives j in the choice set.

$$IND_{in} = \frac{1}{\sum_j f_i(j)} \tag{32}$$

The similarity itself is measured considering the time gap between corresponding departure (DEP) and arrival (ARR) times and the differences in perceived journey time (PJT) and price.

$$f_i(j) = \left(1 - \frac{x_i(j)}{s_x}\right) \cdot \left(1 - \gamma \cdot min\left\{1, \frac{s_z \cdot |y_i(j)| + s_y \cdot |z_i(j)|}{s_y \cdot s_z}\right\}\right) \tag{33}$$

where $x_i(j) = \frac{|DEP(i)-DEP(j)|+|ARR(i)-ARR(j)|}{2}$,

$y_i(j) = PJT(j) - PJT(i)$, and

$z_i(j) = price(j) - price(i)$.

s_x, s_y and s_z set the range of influence of $x_i(j)$, $y_i(j)$ and $z_i(j)$ respectively. s_y and s_z depend on the sign of $y_i(j)$ and $z_i(j)$ in order to model the asymmetry between connections. If there is a difference in terms of perceived journey time, the superior connection will exert a stronger influence on the inferior one, the same applies for the price.

Weis (2006) and van Eggermond *et al.* (2007) applied this measure. Weis (2006) examined ground-based public transport and had no price data available. He found, that similarities between alternatives have a negative influence on their choice probabilities, which complies with the assumptions of the IAP model. Contrarily, van Eggermond *et al.* (2007), who analysed air transport choice and had price data available, determined a positive influence of similarities. This is in accordance with earlier findings by Hoogendoorn-Lanser and Bovy (2007) or Frejinger and Bierlaire (2007), who showed that similarities can have a positive influence on the utility of ground-based transport alternatives. Possible interpretations are that a positive influence can be derived from the possibility to switch routes or connections while the passenger is traveling or to a strong preference for certain alternative attributes that are also present in the chosen alternative such as a specific departure time, travel time or fare.

6.6 The concept of dominance

The concept of dominance has been lately introduced by Cascetta *et al.* (2007) for the context of residential location choice. The basic assumption is that an alternative is less likely to be taken into account if it is dominated by other alternatives. Thus, a dominance factor DF_{in} is calculated for each alternative i, indicating the number of alternatives dominating i. Analogous to the IAP model DF_{in} can than be included in the utility function. On the other hand it can be used for choice set generation.

According to Cascetta *et al.* (2007) an alternative j dominates alternative i, if the utility of all characteristics of j is higher than (or equal to) that of the equivalent characteristic of i. In addition, even stronger dominance rules can be defined by the modeller with the help of thresholds

or specific similarity factors. Cascetta *et al.* (2007) cite two specifications of dominance measures for their problem of residential location choice that use distance as impedance measure.

Cascetta and Papola (2005) assume in their measure for destination choice that alternative j dominates alternative i if the attractiveness of j is greater than that of i while at the same time the generalised costs c_{oj} of getting from origin o to destination j are smaller than those of getting from o to i.

The second dominance measure originates from Stouffer (1960) and refers to the concept of intervening opportunities. In order to dominate i, destination j has to fulfil the conditions formulated by Cascetta and Papola (2005) and in addition being situated on the path from origin o to destination i. In this case j is an intervening opportunity on the path to i.

6.7   The Sequence Alignment method

Joh *et al.* (2001) used the Sequence Alignment Method (SAM) to examine the similarity between alternatives, that compromise multiple characteristics, which themselves have a multivariate description. A transport example are activity patterns. Activity patterns consist of multiple activities, that each have several properties such as type, mode, location, duration. The SAM employs the concept of biological distance rather than geometrical distance. Biological distance is defined as the smallest number of attribute changes (mutations) that is necessary to equalise two alternatives. With the help of this measure not only the types of attributes are considered but also their sequential order. This facts makes the SAM extremely valuable for activity pattern analysis. It is very flexible and allows to determine a simple measure of similarity even for alternatives with different types of attributes and complex interdependencies.

Joh *et al.* (2001) did not apply their similarity measure to discrete choice modelling but used it for the classification of activity patterns and for goodness-of-fit measures in activity based modelling. However, it is an promising approach to capture similarities of multi-dimensional alternatives and can be used in discrete choice modelling through the IAP model.

6.8   Dependencies between decision-makers

The focus of the work presented by Mohammadian *et al.* (2005) was set on the introduction of spatial dependencies between decision-makers instead of alternatives. They developed a Mixed Logit model for new housing projects that accounts for taste heterogeneity and correlations between alternatives. In addition, a spatial dependency parameter ρ is introduced into the systematic part of the utility function to account for spatial correlation between the decision-

makers.

$$U_{in} = V'_{in} + \sum_{s=1}^{S} \rho_{nsi} y_{si} + \varepsilon_{in} \tag{34}$$

where $s = 1, ..., S$ are the decision-makers who's choice influences the choice of decision-maker n while evaluating alternative i and y_{in} is equal to one if decision-maker s has chosen alternative i and zero otherwise. The spatial parameter ρ is a matrix of coefficients representing the influence that the choice of one decision-maker has on another decision-maker while he chooses alternative i. According to Mohammadian *et al.* (2005) it is defined as

$$\rho_{nsi} = \lambda e^{-\frac{D_{ns}}{\gamma}} \tag{35}$$

with D_{ns} being the spatial distance separating decision-maker n and s, and λ and γ being parameters to be estimated.

Dugundji and Walker (2005) also focussed on the explicit account for dependencies between decision-makers instead of alternatives. They employed a so-called Field Effect Variable in the deterministic part of the utility function. This variable represents the dependency of a decision-maker's choice on the overall share of connected decision-makers that choose the alternative in question. However, instead of capturing only spatial dependencies, they suggest a network structure to represent any kind of dependencies between decision-makers, especially social ones. In the dependency network, each decision-makers is symbolised by a node and his or her dependencies by links. Correlations between alternatives and taste heterogeneities in this model have been captured by a CNL model.

7   Evaluation of the different approaches and conclusions

Today the significance of an appropriate representation of similarities between discrete choice alternatives is undoubted and a lot of research effort has been dedicated to this problem. However, no completely satisfactory solution has yet been developed for the context of transportation modelling. Transportation choice problems such as route choice or destination choice are characterised by large sets of alternatives with often overlapping characteristics. Hence, models are needed that are able to handle large choice sets and do not require too much effort for computation, specification and identification and are thus applicable to practical problems. On the other hand similarities between alternatives can be very complex. They can be related to different attributes and have different levels of influence on the utility a decision-maker receives from a specific alternative. Thus, suitable approaches have to be flexible and able to accommodate various and complex similarity structures.

This is especially true for a combined route mode and destination choice model. As the studies cited above show, similarities exist for each individual step of the traditional four step approach. In addition, the model steps themselves are interrelated. Destinations can for example be situated on the same route or can be reached with the same transport mode. Consequently, only a simultaneous consideration of route, mode and destination choice would allow a realistic representation of transport decisions. However, such a model is highly complex, even without accounting for similarities. Thus, computational efficiency is even more important than for models comprising only one model step. On the other hand, also the similarities between alternatives in such a model tend to comprehend multiple levels and dimensions and need to be thoroughly defined. Hence, none of the similarity measures presented in this paper is offhand suitable for a combined route, mode and destination choice model. Yet, some of them are promising starting points.

The family of GEV models can provide the framework for the combined route, mode and destination choice model. Due to the closed form formulation of these models, especially the Nested Logit model will be useful to capture parts of the similarities for example of mode choice alternatives. In addition these models can be combined with other similarity measures in a straight-forward way. An application of the Cross-Nested Logit model, however, is probably not reasonable. Though it is able to depict nearly all kind of correlation structures, its formulation, identification and estimation are already very costly for one level choice situation. This makes it not useful for a combined model with a large set of alternatives and multi-level correlations.

The same applies for the Probit model and most of the Mixed Logit formulations presented in section 5. They represent similarities and correlations between alternatives of the choice set very well and are applicable to any choice situation as demonstrated by various examples.

Furthermore they improve the overall model fit significantly and much more than any of the similarity factor models. However, the effort needed for the specification, identification and estimation of the model, also including simulation methods for estimation as well as application, makes it most likely that these models can not be used for the problem at hand. This goes notwithstanding that some of the approaches presented here, especially the subnetwork model by Frejinger and Bierlaire (2007), constitute interesting approaches and should be further investigated in this context.

The general idea of the models described in section 6 to apply similarities factors in the deterministic part of the utility function is very appealing because of its simplicity and elegance. Instead of structuring the choice set a priori and taking the chance of misleading assumptions about correlations, only the type of similarities is specified. This type accounts for the individual characteristics of the alternatives in the choice set and imposes a value to the impact of specific interdependencies. Practical applications of the models described here demonstrated, that the IIA property has been well accounted for and the models could be estimated with relatively low computational costs even for large sets of alternatives.

However, these models also suffer from some shortcomings. For example they do not take into account taste heterogeneity. But more important they are designed with respect to a specific choice context and usually miss some aspects of the correlation between alternatives. While similarity factors for some choice situations have been extensively investigated and appropriate factors, such as the Path Size factor for private transport route choice, have been established similarities in other choice situations have hardly been tackled. Particularly public transport connection choice and destination choice need further investigation to derive a combined route, mode and destination choice model.

Another research issue concerns the question wether similarities have positive or negative influence on the utility. Theory suggests a negative influence since the probability of being perceived as a distinct alternative is decreased. However, recent studies demonstrated that this can be superimposed by positive evaluations of similarities derived for example from the possibility of switching between alternatives. This question requires further investigation.

References

Abay, G. (1999) *Nachfrageabschätzung Swissmetro: Eine Stated-Preference Analyse*, F1 edn., Berichte des Nationalen Forschungsprogrammes 41: Verkehr und Umwelt, Eidgenössische Drucksachen- und Materialzentrale, Bern.

Abbe, E., M. Bierlaire and T. Toledo (2007) Normalization and correlation of cross-nested logit models, *Transportation Research Part B: Methodological*, **41** (7) 795–808.

Bekhor, S., M. E. Ben-Akiva and M. S. Ramming (2001) Adaptation of logit kernel to route choice situation, in TRB (ed.) *the 80th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2001.

Ben-Akiva, M. E. (1973) Structure of passenger travel demand models, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.

Ben-Akiva, M. E. (1974) Structure of passenger travel demand models, *Transportation Research Record*, **526**.

Ben-Akiva, M. E. and M. Bierlaire (1999) Discrete choice methods and their applications to short-term travel decisions, in R. Hall (ed.) *Handbook of Transportation Science*, 5–34, Kluwer, Dordrecht.

Ben-Akiva, M. E. and D. Bolduc (1996) Multinomial probit with a logit kernel and a general parametric specification of the covariance structure, in ICS (ed.) *the 3rd Invitational Choice Symposium*.

Ben-Akiva, M. E. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.

Bierlaire, M. (2002) The Network GEV model, in STRC (ed.) *the 2nd Swiss Transport Research Conference (STRC)*, Ascona, Mar. 2002, `http://www.strc.ch/2002.html`.

Bierlaire, M., K. W. Axhausen and G. Abay (2001) The acceptance of modal innovation : The case of swissmetro, in STRC (ed.) *the 1st Swiss Transport Research Conference (STRC)*, Ascona, Mar. 2001, `http://www.strc.ch/2001.html`.

Borgers, A. W. J. and H. J. P. Timmermans (1987) Choice model specification, substitution and spatial structure effects: A simulation experiment, *Regional Science and Urban Econimics*, **17**, 29–47.

Cascetta, E., A. Nuzzola, F. Russo and A. Vitetta (1996) A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, in J. B. Lesort (ed.) *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, 697–711, Pergamon, Oxford.

Cascetta, E., F. Paliara and K. W. Axhausen (2007) The use of dominance variables in choice set generation, in WCTRS (ed.) *the 11th World Conference on Transportation Research*, Berkeley, Jun. 2007.

Cascetta, E. and A. Papola (2003) A joint mode-transit service choice model incorporating the effect of regional transport service timetables, *Transportation Research Part B: Methodological*, **37** (7) 595–614.

Cascetta, E. and A. Papola (2005) Dominance among alternatives in random utility models: A general framework and an application to destination choice, in ETC (ed.) *European Transport Conference*, Strasbourg, Oct. 2005.

Cascetta, E., F. Russo, F. A. Viola and A. Vitetta (2002) A model of route perception in urban road networks, *Transportation Research Part B: Methodological*, **36** (7) 577–592.

Daganzo, C. F. (1979) *Multinomial Probit: The theory and its application to demand forecasting*, Academic Press, New York.

Daganzo, C. F. and Y. Sheffi (1977) On stochastic models of traffic assignment, *Transportation Science*, **11** (3) 253–274.

Debreu, G. (1960) Review of R.D. Luce individual choice behavior, *American Economic Review*, **50**, 186–188.

Dugundji, E. R. and J. L. Walker (2005) Discrete choice with social and spatial network interdepencies: An empirical example using mixed gev models with field and "panel" effects, in TRB (ed.) *the 84th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2005.

Fotheringham, A. S. (1983) A new set of spatial interaction models: The theory of competing destinations, *Environment and Planning A*, **15**, 15–36.

Fotheringham, A. S. (1988) Consumer store choice and choice set definition, *Marketing Science*, **7** (3) 299–310.

Frejinger, E. and M. Bierlaire (2007) Capturing correlation with subnetworks in route choice models, *Transportation Research Part B: Methodological*, **41** (3) 363–378.

Friedrich, M., I. Hofsäss and S. Wekeck (2001) Timetable-based transit assignment using branch and bound techniques, *Transportation Research Record*, **1752**, 100–107.

Gower, J. C. (1985) Measures of similarity, dissimilarity, and distance, in S. Kotz, N. L. Johnson and C. B. Read (eds.) *Encyclopedia of Statistical Sciences*, vol. 5, 397–405, John Wiley & Sons, New York.

Guo, J. Y. and C. R. Bhat (2005) Operationalizing the concept of neighborhood: Application to residential location choice analysis, in TRB (ed.) *the 84th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2005.

Hess, S., M. Bierlaire and J. W. Polak (2005) Capturing taste heterogeneity and correlation structures with Mixed GEV models, in TRB (ed.) *the 84th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2005.

Hoogendoorn-Lanser, S. and P. H. L. Bovy (2007) Modeling overlap in multi-modal route choice by inclusion of trip part specific path size factors, in TRB (ed.) *the 86th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2007.

Hoogendoorn-Lanser, S., R. van Nes and P. H. L. Bovy (2005) Path size and overlap in multimodal transport networks, in H. S. Mahmassani (ed.) *Flow, Dynamics and Human Interaction - Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, 63–83, Elsevier, Oxford.

Joh, C.-H., T. A. Arentze and H. J. P. Timmermans (2001) A position-sensitive sequence alignment method illustrated for space-time activity-diary data, *Environment and Planning A*, **33** (2) 313–338.

Kitamura, R. (1984) Incorporating trip chaining into analysis of destination choice, *Transportation Research Part B: Methodological*, **18** (1) 67–81.

Koppelman, F. S. and C.-H. Wen (2000) The paired combinatorial logit model: Properties, estimation and application, *Transportation Research Part B: Methodological*, **34** (2) 75–89.

McFadden, D. (1974) Conditional logit analysis of qualitative choice-behaviour, in P. Zarembka (ed.) *Frontiers in Econometrics*, Academic Press, New York.

McFadden, D. (1978) Modeling the choice of residential location, in A. Karlqvist (ed.) *Spatial Interaction Theory and Residential Location*, 75–96, North-Holland, Amsterdam.

Mohammadian, A. K., M. Haider and P. S. Kanaroglou (2005) Incorporating spatial dependencies in random parameter discrete choice models, in TRB (ed.) *the 84th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2005.

Papola, A. (2004) Some development on the cross-nested logit model, *Transportation Research Part B: Methodological*, **38** (9) 833–851.

Prato, C. G. and S. Bekhor (2007) Modeling route choice behavior: How relevant is the choice set composition?, in TRB (ed.) *the 86th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan. 2007.

Ramming, M. S. (2002) Network knowledge and route choice, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.

Small, K. A. (1987) A discrete choice model for ordered alternatives, *Econometrica*, **55** (2) 409–424.

Stouffer, S. A. (1960) Intervening opportunities and competing migrants, *Journal of Regional Science*, **2**, 1–26.

van Eggermond, M., N. Schüssler and K. W. Axhausen (2007) Accounting for similarities in air transport route choice, *Working Paper*, **XXX**, IVT, ETH Zurich, Zurich, `https://www.ivt.ethz.ch/vpl/publications/reports/index/edit/abxxx.pdf`.

Vovsha, P. and S. Bekhor (1998) The link-nested logit model of route choice: Overcoming the route overlapping problem, *Transportation Research Record*, **1645**, 133–142.

Vrtic, M. (2003) Simultanes Routen- und Verkehrsmittelwahlmodell, Ph.D. Thesis, Technical University Dresden, Dresden.

Walker, J. L. (2001) Extended discrete choice models: Integrated framework, flexible error structures, and latent variables, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.

Walker, J. L. (2002) The mixed logit (or logit kernel) model: Dispelling misconceptions of identification, *Transportation Research Record*, **1805**, 86–98.

Walker, J. L., M. E. Ben-Akiva and D. Bolduc (forthcoming) Identification of parameters in normal error component logit-mixture (NECLM) models, *Journal of Applied Econometrics*.

Weis, C. (2006) Routenwahl im ÖV, Master Thesis, IVT, ETH Zurich, Zurich.

Wen, C.-H. and F. S. Koppelman (2001) The generalized nested logit model, *Transportation Research Part B: Methodological*, **35** (7) 627–641.

Yai, T., S. Iwakura and S. Morichi (1997) Multinomial probit with structured covariance for route choice behaviour, *Transportation Research Part B: Methodological*, **31** (3) 195–208.