CLOSED FORM DISCRETE CHOICE MODELS

Frank S. Koppelman

And

Vaneet Sethi

Department of Civil Engineering
Northwestern University
Evanston, Illinois, 60208, U.S.A.

INTRODUCTION

Random utility maximization discrete choice models are widely used in transportation and other fields to represent the choice of one among a set of mutually exclusive alternatives. The decision maker, in each case, is assumed to choose the alternative with the highest utility to him/her. The utility to the decision maker of each alternative is not completely known by the modeler; thus, the modeler represents the utility by a deterministic portion which is a function of the attributes of the alternative and the characteristics of the decision-maker and an additive random component which represents unknown and/or unobservable components of the decision maker's utility function.

Early development of choice models was based on the assumption that the error terms were multivariate normal or independently and identically Type I extreme value (gumbel) distributed (Johnson and Kotz, 1970). The multivariate normal assumption leads to the multinomial probit (MNP) model (Daganzo, 1979); the independent and identical gumbel assumption leads to the multinomial logit (MNL) model (McFadden, 1973). The probit model allows complete flexibility in the variance-covariance structure of the error terms but it's use requires numerical integration of a multi-dimensional normal distribution. The multinomial logit probabilities can be evaluated directly but the assumption that the error terms are independently and identically distributed across alternatives and cases (individuals, households or choice repetitions) places important limitations on the competitive relationships among the alternatives. Developments in the structure of discrete choice models have been directed at either reducing the computational burden associated with the multinomial probit model (McFadden, 1989; Hajivassiliou and McFadden, 1990; Börsch-Supan and Hajivassiliou, 1992; Keane, 1994) or increasing the flexibility of extreme value models.

Two approaches have been taken to enhance the flexibility of the MNL model. One approach, the development of open form discrete choice models is discussed by Bhat in another chapter of this handbook. This chapter describes the development of closed form models which relax the assumption of independent and identically distributed random error terms in the multinomial logit model to provide a more realistic representation of choice probabilities. The rest of this chapter is structured as follows. The next section reviews the properties of the multinomial logit model as a reference point for the development and interpretation for other closed form models. The next two sections describe models that relax the assumption of independence of error terms across

alternatives and models that relax the assumption of equality of error terms across cases. The final section suggests directions for further development of these models.

MULTINOMIAL LOGIT MODEL

The multinomial logit (MNL) model is derived through the application of utility maximization concepts to a set of alternatives from which one, the alternative with maximum utility, is chosen. The modeler assumes the utility of an alternative i to an individual q, $U_{i,q}$, includes a deterministic component, $V_{i,q}$, and an additive random component, $e_{i,q}$; that is,

$$U_{i,q} = V_{i,q} + e_{i,q} \tag{1}$$

The deterministic component of the utility function, which is commonly specified as linear in parameters, includes variables which represent the attributes of the alternative, the decision context and the characteristics of the traveler or decision maker. The linearity of the utility function can be overcome by prior transformation of variables, quadratic forms, spline functions (line segment approximations) or estimation with special purpose software.

Assuming that the random component, which represents errors in the modeler's ability to represent all of the elements which influence the utility of an alternative to an individual, is independently and identically gumbel distributed across cases and alternatives leads to the multinomial logit model:

$$P_{i,q} = \frac{e^{V_{i,q}}}{\sum_{i'=1}^{J} e^{V_{i',q}}} \tag{2}$$

where $P_{i,q}$ is the probability that alternative i is chosen by individual q,
$e^{(\ )}$ is the exponential function,
$V_{i,q}$ is the deterministic component of the utility of alternative i for individual q, and
J is the number of alternatives

The closed form of the MNL model makes it straightforward to estimate, interpret and use. As a result, the MNL model has been used in a wide variety of travel and travel related choice contexts including mode, destination, car ownership and residential location as well as choices in non-travel contexts.

However, the assumptions that the error terms are distributed independently and identically across cases and alternatives are likely to be violated in many choice contexts. The development of alternative structural model forms has been directed toward the relaxation of these assumptions. Attempts to relax the equality of the variance of the error distribution over alternatives have been based on the MNP or other models which require numerical evaluation of multi-dimensional integrals. The relaxation of the assumptions of independence across alternatives and equal variance across cases can be achieved through the use of closed form models as well as through the adoption of models which require numerical integration. Before examining these closed form relaxations of the MNL, we review the impact of the independent and identically distributed error assumptions.

Independence of Errors Across Alternatives
The independence and equal variance of error terms across alternatives leads to the property of independence of irrelevant alternatives (IIA) which states that the relative probability of choosing any pair of alternatives is

independent of the presence or attributes of any other alternatives. This is illustrated by the equality of cross-elasticities of the probabilities of all other alternatives, i' in response to a change in any alternative, i,

$$\boldsymbol{h}_{X_{i,k}}^{P_{i'}} = -P_i \boldsymbol{b}_k X_{i,k} \qquad (3)$$

where $\boldsymbol{h}_{X_{i,k}}^{P_{i'}}$ represents the cross-elasticity of the probability of alternative i' to a change in the kth variable describing an attribute of alternative i, $X_{i,k}$, and

$\boldsymbol{b}_k$ is the parameter associated with $X_{i,k}$.

Thus, the assumption of independent (uncorrelated) errors across alternatives results in equal proportional substitution between alternatives.

Equality of Error Variance across Cases

The assumption of equal variance of the error components of the utility functions across cases as well as alternatives may be inappropriate in a variety of contexts. For example, in the case of mode or path choice in which the travel distance varies widely across cases, the variance of the unobserved components of utility is likely to increase with distance. This and other similar cases can be addressed by formulating and estimating a relationship between the error variance parameter or scale of the utility and variables, such as distance, which describe the choice context and/or the characteristics of the decision maker.

The rest of this chapter describes approaches which have been taken to relax the assumptions of independence among alternatives and equal variance over cases within the context of closed form models based on extreme value error distributions.

## RELAXATION OF THE INDEPENDENCE OF ERRORS ACROSS ALTERNATIVES

The nested logit model was the first closed form alternative to the MNL and has been the most widely used alternative over the last two decades. During the same period, two general approaches were proposed to further relax the restrictions associated with the MNL. The generalized extreme value (GEV) family of models (McFadden, 1978) provides a theoretical basis to relax the independence assumption and cross-elasticity restrictions of the NL and MNL models within the utility maximization framework. Specific GEV models have been shown to be statistically superior to both the MNL and NL models in transportation applications. The universal logit model family relaxes the MNL cross-elasticity properties by including attributes of other alternatives in the deterministic portion of the utility function of some or all of the alternatives in the choice set. This section discusses the properties of these groups of models.

The Nested Logit Model

The nested logit (NL) model allows dependence or correlation between the utilities of alternatives in common groups (Williams, 1977; Daly and Zachary, 1978; McFadden, 1978). Derivation of the NL model is based on the same assumptions as the MNL model except that correlation of error terms is assumed to exist among pre-defined groups of alternatives. Such error correlations arise if an unobserved factor influences the utility of all members of the group. The NL model can be written as the product of a series of MNL choice models defining each level in a tree structure. For example, the tree depicted in Figure 1 includes three levels I, J, and K. Each elemental alternative is represented at the Ith level. The number of alternatives in each nest at each nest level varies; in particular, some nests, such as 1, 11 and 22, may have only a single member. In this case the structural parameter is identically equal to one and the probability of the alternative, given the nest, is one.

The values of the structural parameters, q_{jk} and q_k, indicate the substitution or cross-elasticity relationship between alternatives in the nest. To be consistent with utility maximization, the structural parameters at the highest level and the ratios of the structural parameters at each lower nest are bounded by zero and one. The estimated parameters at each node represent the ratio between the structural parameter at that node and at the next higher node in the tree. A value of one for any ratio of structural parameters implies that the alternatives in that nest are uncorrelated and can be directly connected to the next higher node. If all the structural parameter ratios equal one, all the alternatives can be directly linked to the root of the tree; that is, the structure collapses to the MNL.

Figure 1: A General Tree Notation for Three Level Nested Logit Structure
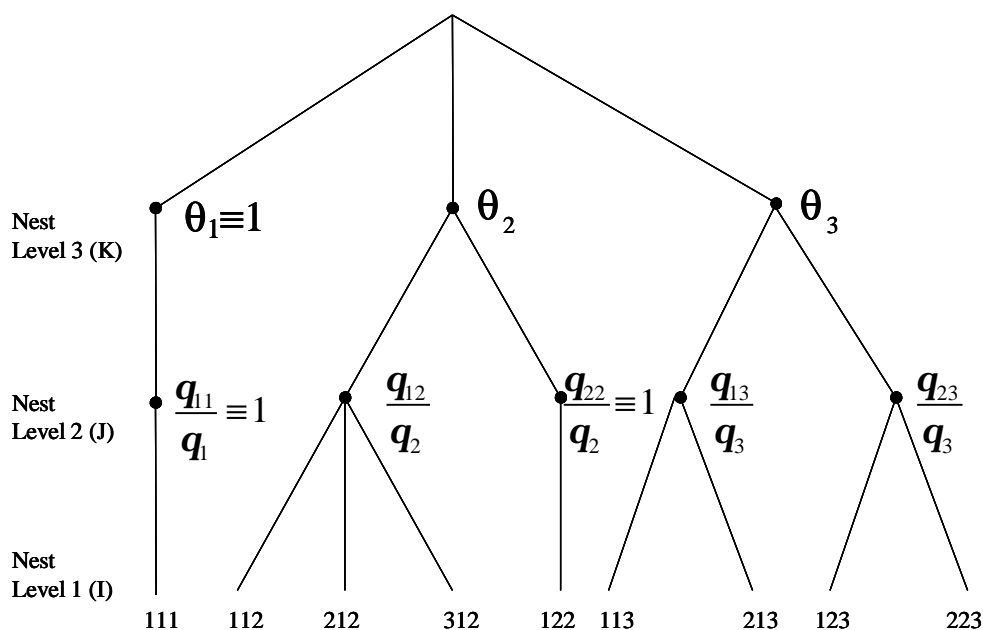
The probability associated with each elemental alternative, is given by the product of the probabilities for each choice level between the root of the tree and that alternative; that is,

$$P_{ijk} = P_{i/jk} \times P_{j/k} \times P_k$$

$$P_{ijk} = \frac{e^{\left(\frac{V_{ijk}}{q_{jk}}\right)}}{\sum\limits_{i'} e^{\left(\frac{V_{i'jk}}{q_{jk}}\right)}} \times \frac{e^{\left(\frac{q_{jk}}{q_k} \times \Gamma_{jk}\right)}}{\sum\limits_{j'} e^{\left(\frac{q_{j'k}}{q_k} \times \Gamma_{j'k}\right)}} \times \frac{e^{(q_k \Gamma_k)}}{\sum\limits_{k'} e^{(q_{k'} \times \Gamma_{k'})}} \tag{4}$$

$$\Gamma_{jk} = \ln \sum\limits_{i'} e^{\frac{V_{i'jk}}{q_{jk}}} ; \quad \Gamma_k = \ln \sum\limits_{j'} e^{\frac{q_{j'k}}{q_k} \times \Gamma_{j'k}}$$

Thus, each alternative in the nest can be connected to the next higher node in the tree eliminating the nest. As any structural parameter decreases, the similarity between alternatives in the nest increases and the sensitivity of an alternative to changes in other alternatives in the nest also increases; that is, $h^{P_{i'}}_{X_{i,k}} = -P_i b_k X_{i,k} q_{i,i'}$ where $q_{i,i'}$ is the structural parameter for the lowest nest which includes both alternatives.

The assignment of alternatives to positions in the tree and the overall structure of the tree are subject to the discretion of the modeler who can impose an *a priori* structure or search over some or all of the possible nesting structures. A practical problem with the NL model is the large number of possible nesting structures for even moderate numbers of alternatives. For example, five alternatives can result in 250 distinct structures (65 two-level structures, 125 three level structures and 60 four level structures). On the other hand disallowing some structures *a priori*, based on reasonableness criteria, can result in the loss of insight from the empirical testing.

The NL model, by allowing correlation among subsets of utility functions, alleviates the IIA problem of the MNL, but only in part. It retains the restrictions that alternatives in a common nest have equal cross-elasticities and alternatives not in a common nest have cross-elasticities as for the MNL.

Generalized Extreme Value Models

The generalized extreme value (GEV) family of models (McFadden, 1978) relaxes the error independence and substitution relationships among alternatives. The GEV family includes all closed form utility maximization formulations based on the extreme value error distribution with equal variance across alternatives. GEV models can be generated from any function of Y_i for each alternative, i,

$$G(Y_1, Y_2, \ldots, Y_J), \ Y_1, Y_2, \ldots, Y_J \geq 0 \tag{5}$$

which is non-negative, homogeneous , goes to infinity with each Y_i and has odd (even) order partial derivatives which are non-negative (non-positive). The probability equation of such a model, for μ equal to one and the transformation, $Y_i = \exp(V_i)$, to ensure positive Y_i is

$$P_i = \frac{e^{V_i} G_i(e^{V_1}, e^{V_2}, \ldots, e^{V_J})}{G(e^{V_1}, e^{V_2}, \ldots, e^{V_J})} \tag{6}$$

where $G_i(\ )$ is the first derivative of G with respect to Y_i, and

V_i represents the observable component of the utility for each alternative.

Both the multinomial logit and nested logit models belong to the GEV family. GEV models can be grouped into two classes. One class, models that are flexible with respect to cross-elasticities between alternatives, includes the paired combinatorial logit (PCL) model (Chu, 1989, Koppelman and Wen, 1998), the cross-nested logit (CNL) model (Vovsha, 1997), ~~and~~ the generalized nested logit (GNL) model (Wen and Koppelman, 2000), the generalized MNL (GenMNL) model (Swait, 2000), and the fuzzy nested logit (FNL) model (Vovsha, 1999). Each of these models allows differential correlation or rates of substitution between pairs of alternatives. The second class includes models, which impose a specific structural relationships on cross-elasticities between alternatives, *a priori*. Models in this class are the ordered generalized extreme value model (Small, 1987), which allows differential correlation among alternatives based on their proximity in an ordered set; the principles of differentiation (PD) model (Bresnahan *et. al.*, 1997); and the cross-correlated logit (CCL) model (Williams, 1977), which allow differences in correlation between alternatives along distinct dimensions of choice.

The models in the first class are differentiated by the way in which they incorporate pairwise correlation/substitution into the model. The PCL model assigns equal portions of each alternative to one nest with each other alternative. The total probability of choosing an alternative is the sum over pairs of alternatives of the unobserved probability of the pair times the probability of the alternative given choice of that pair.

$$P_i = \sum_{j \neq i} P_{i/ij} \times P_{ij}$$

$$= \sum_{j \neq i} \left[\frac{\left(a\ e^{V_i}\right)^{\frac{1}{q_{ij}}}}{\left(a\ e^{V_i}\right)^{\frac{1}{q_{ij}}} + \left(a\ e^{V_j}\right)^{\frac{1}{q_{ij}}}} \times \frac{\left(\left(a\ e^{V_i}\right)^{\frac{1}{q_{ij}}} + \left(a\ e^{V_j}\right)^{\frac{1}{q_{ij}}}\right)^{q_{ij}}}{\sum_{k=1}^{J-1} \sum_{m=k+1}^{J} \left(\left(a\ e^{V_k}\right)^{\frac{1}{q_{km}}} + \left(a\ e^{V_m}\right)^{\frac{1}{q_{km}}}\right)^{q_{km}}} \right] \quad (7)$$

where $P_{i/ij}$ is the conditional probability of choosing alternative i given the choice of pair ij,
   P_{ij} is the marginal probability for the alternative pair ij,
   V_i is the observable portion of the utility for alternative i,
   J is the number of alternatives in the choice set and
   a is the fraction of i assigned to nest ij and is equal to 1/(J-1) for all alternatives and nests.

The summation includes all pairs of alternatives in the choice set of *J* alternatives, and q_{ij} is the structural parameter associated with alternative pair *ij*. This formulation allows different cross-elasticities to be estimated for each pair of alternatives. However, the equal allocation of each alternative to a nest with each other alternative limits its maximum implied correlation with each other alternative to 1/(J-1) which similarly limits the maximum cross-elasticity for the pair. This limitation may be serious in cases with more than a few alternatives.

The CNL model allows different proportions of each alternative to be assigned to nests selected by the modeler with each nest having the same structural parameter, the probability of each alternative being:

$$P_i = \sum_m P_{i/m} \times P_m = \sum_m \left[\frac{\left(a_{im}\ e^{V_i}\right)^{\frac{1}{q}}}{\sum_{j \in N_m} \left(a_{jm}\ e^{V_j}\right)^{\frac{1}{q}}} \times \frac{\left(\sum_{j \in N_m} \left(a_{jm}\ e^{V_j}\right)^{\frac{1}{q}}\right)^{q}}{\sum_m \left(\sum_{j \in N_m} \left(a_{jm}\ e^{V_j}\right)^{\frac{1}{q}}\right)^{q}} \right] \qquad (8)$$

where V_i is the observable portion of the utility for alternative i,

N_m is the set of all alternatives included in nest m,

q is the similarity parameter for all nests, $0 < q \le 1$,

a_{im} is the portion of alternative i assigned to nest m and must satisfy the conditions that

$$\sum_m a_{im} = 1, \ \forall i \ \text{and} \ a_{im} > 0 \ \forall i,m.$$

The implied correlation and substitution between alternatives is determined by the fractions of each alternative included in one or more common nests. The selection of the number of nests and the assignment of alternatives to each nest are left to the judgement of the modeler. The constraint of equal logsum parameters for all nests limits the maximum cross-elasticity associated with each pair of alternatives.

The GNL model, which combines the flexibility of the PCL model (different structural parameters for each nest) with that of the CNL model (different proportions of each alternative assigned to each nest), enables very flexible correlation/substitution patterns. The choice probabilities for the GNL model are given by:

$$P_i = \sum_m P_{i/m} \times P_m = \sum_m \left[\frac{\left(a_{im}\ e^{V_i}\right)^{\frac{1}{q_m}}}{\sum_{j \in N_m} \left(a_{jm}\ e^{V_j}\right)^{\frac{1}{q_m}}} \times \frac{\left(\sum_{j \in N_m} \left(a_{jm}\ e^{V_j}\right)^{\frac{1}{q_m}}\right)^{q_m}}{\sum_{m'} \left(\sum_{j \in N_{m'}} \left(a_{jm'}\ e^{V_j}\right)^{\frac{1}{q_{m'}}}\right)^{q_{m'}}} \right] \qquad (9)$$

Swait (2000) proposed the Generalized MNL model, to simultaneously evaluate choice and choice set generation. The GenMNL model is identical to the GNL except that the allocation parameters are constrained to be equal. Vovsha (1999) reports development and application of the Fuzzy Nested Logit model, which is identical to the GNL, except that it allows multiple levels of nesting. While this model is technically different from the GNL model, we believe that the GNL can closely approximate the FNL model.

The PCL and CNL are special cases of the GNL model with allocation parameters constrained to be equal $(a_{jm} = a, \forall j,m)$ for the PCL model and the structural parameters constrained to be equal $(q_m = q, \forall m)$ for the CNL model.

The differences among these models and the MNL and NL models are illustrated by differences in their cross-elasticities (Table 1). In each case, the cross elasticity is a function of the probability of the alternative which is changed, the value of the variable which is changed and the parameter associated with that variable. The differences among models are represented by additions to the probability term. These additions are functions of the structural parameter(s) for nests that include the changed alternative and the alternative for which the elasticity is formulated and, for the CNL and GNL models, increases with the allocation parameters, embedded in the conditional probabilities, for these alternatives.

The CNL, GNL, GenMNL and FNL models all require the analyst to choose among a large set of possible nesting structures which, at the limit, includes single alternative nests, and nests with all possible combinations of alternatives. Swait (2000) explains these alternative nests in terms of the choice sets which might feasibly be generated and compares models with different nesting structures. The search requirement, which is similar to the NL search problem, places responsibility on the analyst to explore and select among many structural alternatives. The PCL model, that strictly limits the assignment of alternatives to nests, does not share this problem. One approach to the search problem is to implement a paired GNL model and to use the results to identify groups of alternatives that might be included in a common nest.

The models in the second class assume specific structural relationships among alternatives. The ordered generalized extreme value (OGEV) model (Small, 1987) allows correlation and, thus, the substitution between alternatives in an ordered choice set to increase with their proximity in that order. Each alternative is a member of nests with one or more adjacent alternatives. The general OGEV model allows different levels of substitution by changing the number of adjacent alternatives in each nest, the allocation weights of each alternative to each nest and the structural parameters for each nest. The choice probabilities for the general OGEV model is given by

$$P_i = \sum_{m=i}^{i+L} P_{i/m} \times P_m = \sum_{m=i}^{i+L} \left[\frac{\left(a_{m-i}\, e^{\frac{V_i}{q_m}} \right)}{\sum_{j \in N_m} \left(a_{m-j}\, e^{\frac{V_j}{q_m}} \right)} \times \frac{\left(\sum_{j \in N_m} \left(a_{m-j}\, e^{\frac{V_j}{q_m}} \right) \right)^{q_m}}{\sum_{s=1}^{J+L} \left(\sum_{j \in B_s} \left(a_{s-j}\, e^{\frac{V_j}{q_s}} \right) \right)^{q_m}} \right] \quad (10)$$

where V_i is the observable portion of the utility for alternative i

      L   is a positive integer that defines the maximum number of contiguous alternatives that can be included in a nest,

      J is the total number of alternatives,

      N_m is the set of all alternatives included in nest m,

      q_m is the similarity parameter that satisfies the condition, $0 < \theta_m \le 1$, and

      a_m is an allocation parameter satisfying the condition that $\sum_{m=0}^{M} a_m = 1$

      $e^{\frac{V_i}{q}}$ is equal to zero for i < 1 and i > J.

The principles of differentiation (PD) model (Bresnahan *et. al.*, 1997) is based on the notion that markets for differentiated products (alternatives) exhibit some form of clustering (nesting) relative to dimensions which characterize some attribute of the product. Under this structure, alternatives that belong to the same cluster compete more closely with each other than with alternatives belonging to other clusters. The PD model defines such clusters along multiple dimensions. The choice probability equations for a PD model with *D* dimensions of differentiation and j_d levels along each dimension, is given by:

$$P_i = \sum_{d \in D} \sum_{j \in d} \left(P_{i|j,d} \times P_{j,d} \right)$$

$$= \sum_{d \in D} \sum_{j \in d} \left(\frac{e^{\frac{V_i}{q_d}}}{\sum_{k \in d} e^{\frac{V_k}{q_d}}} \times \frac{\left[a_d \left(\sum_{k \in d} e^{\frac{V_k}{q_d}} \right)^{q_d} \right]}{\sum_{j \in d'} \sum_{j} a_{d'} \left(\sum_{k'} e^{\frac{V_{k'}}{q_d}} \right)^{q_d}} \right) \tag{11}$$

where V_i is the systematic component of the utility for alternative i,

       q_d is the structural parameter that measures the degree of similarity among products in the same category along dimension d, and

       a_d is the weight for dimension d which can be defined as a function of the structural parameters, q_d, or estimated separately.

The PD structure avoids the need for ordering nests in multi-dimensional choice contexts as is required by use of multi-level NL models and allows cross-elasticities along each dimension; thus, making no *a priori* assumption about the relative importance or similarity of each dimension. The PD structure can be applied to the variety of multi-dimensional choice contexts that occur in transportation modeling such as the joint choice of mode and destination or the three-dimensional choice of residential location, auto ownership and mode to work.

The OGEV and PD models can be shown to be special cases of the GNL model (Wen and Koppelman, 2000).

The cross-correlated logit (CCL) model (Williams, 1977, Williams and Ortuzar, 1982), formulated to account for differences in substitution along two distinct dimensions, is similar in spirit to the PD model. However, the authors of this model adopted a numerical solution technique rather than develop a closed form solution.

The proliferation of GEV models places increased responsibility on the analyst who must select the most appropriate model among the models available. Models that allow increasingly flexible structural relationships among alternatives add to the estimation complexity, computational demands and the time required searching for and selecting a preferred model structure. This task is interrelated with the task of searching for and selecting a preferred utility function specification. In some cases, a relatively simple structure will be adequate to represent the underlying behavior; in others, a relatively complex model structure will be required. The required level of complexity in each case is unlikely to be known *a priori*. However, methods and rules can be developed to guide the search among alternative structures. Nonetheless, analyst judgement and structural exploration is likely to be needed to ensure an appropriate tradeoff between model complexity and ease of estimation, interpretation and use.

Universal (Mother) Logit Models

Another approach to generalization of the MNL model is to include attributes of competing alternatives in the utility function for each alternative (McFadden, 1975). The flexibility of this generalization can be used "to approximate all qualitative choice models with continuous probabilities" giving rise to its description as the "universal" or "mother" logit. Despite its apparent flexibility and use of MNL estimation programs, there are few examples of mother logit models in the literature. This may be due to lack of consistency with utility maximization in some cases, the potential to obtain counter-intuitive elasticities, and the complexity of search for a preferred specification (Ben-Akiva, 1974). Examples of universal logit models, developed and applied to transportation problems, are the Dogit model, the Parameterized Logit Captivity model and the C-Logit model, described in the balance of this section.

The Dogit model (Gaudry and Dagenais, 1978) takes account of cases in which some of the decision-makers are captive or loyal to a particular alternative, to the extent that they choose that alternative independent of its attributes or the attributes of the other alternatives. The Dogit model form is:

$$P_i = \frac{e^{V_i} + p_i \sum_{j=1}^{J} e^{V_j}}{(1 + \sum_{j=1}^{J} p_j) \sum_{j=1}^{J} e^{V_j}} \tag{12}$$

where P_i is the probability of choosing the ith alternative,
  V_i is the observable portion of the utility for alternative i,
  p is a non-negative captivity parameter associated with alternative i, and
  J is the number of alternatives

The Dogit model reduces to the MNL when all p equal zero. Higher values of p imply greater probability that an individual is captive or loyal to alternative i. Relaxation of the IIA constraint of the MNL can be seen in terms of the cross-elasticities of alternatives, represented by

$$h^{P_{i'}}_{X_{i,k}} = b_k X_{i,k} \left[P_i - \frac{p_i}{P_{i'} + p_i} P_i \right] \tag{13}$$

which, unlike the MNL, are different across alternatives. In particular, the elasticity of any alternative decreases with increasing captivity/loyalty of that alternative.

Swait and Ben-Akiva (1987) generalized the Dogit model, which incorporates probabilistic choice set generation, denominated as the Parameterized Logit Captivity (PLC) model. Specifically, in this formulation, the captivity parameter for each alternative *i* is parameterized as a function of independent variables, but restricted to be non-negative. That is

$$p_i = F(g' X_{i,q}) \tag{14}$$

where $X_{i,q}$ is a vector of socio-economic characteristics of decision-maker and attributes of

alternative i, case q,

$\boldsymbol{g}$ is a vector of parameters to be estimated, and

F is a transformation function

The C-Logit model, proposed by Cascetta (1996) in the context of route choice modeling in which alternatives share common elements (links), takes account of such overlapping sections by subtracting a commonality measure from the utility of each alternative to represent the extent of overlap with other alternatives as

$$P_i = \frac{e^{(V_i - CF_i)}}{\sum\limits_{i'=1}^{J} e^{(V_{i'} - CF_{i'})}} \tag{15}$$

where CF_i is the commonality factor of path i with other paths in the choice set

This subtraction results in reduced probability shares for alternatives that share common links; however, the substitution between alternatives is identical for all alternatives, as shown by the cross-elasticity expression; that is the model retains the IIA property of the MNL.

$$\boldsymbol{h}_{X_{i,k}}^{P_{i'}} = -(P_i \boldsymbol{b}_k X_{i,k}) e^{CF_i} \tag{16}$$

Overview of Models which relax the Independence of Errors over Alternatives

The independent and identically distributed error distribution assumption of the MNL model has been widely recognized as producing choice relationships that are likely to be inconsistent with behavioral decision processes; considerable work has been undertaken to develop alternative model forms which relax these constraints. These developments have been based on the generalized extreme value and mother logit models, both of which were proposed by McFadden (1978, 1975). The evolution of these models has been toward increasing relaxation of the independence constraint, which, in turn, relaxes the cross-elasticity property of the MNL. The generalized nested logit model, the most flexible of the models proposed to date, includes the other GEV models as special cases. The advantage of increased flexibility of structure brings with it the need to estimate a larger set of parameters, which may lead to problems of estimation and identification, and imposes an additional search problem for the modeler in his/her consideration of alternative model structures. Thus, there is an important place in the modeler's toolbox for models with restricted structures based on an understanding of the relationships in each choice context.

RELAXATION OF THE EQUALITY OF ERRORS OVER CASES

The models described above assume the equality of error variance-covariance structure across cases; that is, the distribution of information, which is excluded from the utility assessment, is approximately equal across cases. The assumption of variance and/or covariance equality across cases may be inappropriate in a variety of choice situations. Examples include route and mode choice where the error variance is likely to increase with distance, stated preference responses in error variances may increase (decrease) due to respondent fatigue (learning) and differences in choice task complexity that may arise due to the number and/or similarity of the choice alternatives. Similarly, the assumption of covariance homogeneity (or equivalently, equal correlation) may be violated in choice where the degree of substitution between travel modes may vary by trip related attributes (*e.g.*, trip distance) and/or characteristics of the traveler. For example, rail and automobile

may be more substitutable with each other than with air for shorter intercity trips, relative to longer distance trips where air is more likely to be substitutable with rail.

Swait and Adamowicz's (1996) Heteroscedastic Multinomial Logit (HMNL) model allows the random error variances to be non-identical across individuals/cases. The model is motivated by the hypothesis that individuals with the same systematic utility for an alternative may have different abilities to discriminate between the utilities of different alternatives. These differences can be represented in the model as a parameterization of the variance of the random error terms of the utility function. One approach, based on the complexity of a choice situation (E_q) is defined as a function of individual characteristics (*e.g.*, income) and the choice context variables (*e.g.*, number of alternatives, similarity of alternatives, *etc.*). Since the complexity measure is constant across alternatives, the scale factors vary only by case and not by alternative. The choice probabilities for the HMNL model are given as

$$P_{i,q} = \frac{e^{\boldsymbol{m}(E_q) \, V_{i,q}}}{\sum\limits_{i'=1}^{J} e^{\boldsymbol{m}(E_q) \, V_{i',q}}} \tag{17}$$

where $V_{i,q}$ is the systematic component of the utility for alternative *i* case *q* and
$\boldsymbol{m}(E_q)$ is the scale parameter for case t as a function of the complexity measure (E_q) for a
        given choice situation,

This formulation ensures consistency with random utility maximization as shown explicitly by Swait and Adamowicz (1996) and recognized by Ben-Akiva and Lerman (1985, pp. 204-207). The HMNL model retains the same basic properties as the MNL, most notably IIA and uniform cross-elasticities. However, the parameterization of the scale causes the cross-elasticities to differ for each case as follows:

$$\boldsymbol{h}_{X_{i,k}}^{P_{i'}} = -\boldsymbol{m}(E_q) \times (P_i \boldsymbol{b}_k X_{i,k}) \tag{18}$$

The Covariance Heterogeneous Nested Logit (COVNL) model (Bhat, 1997), formulated as an extension of the NL model allows heterogeneity across cases in the covariance of nested alternatives. The COVNL model accommodates covariance (or equivalently, correlation) heterogeneity across cases by parameterizing the structural parameter(s) as function(s) of individual and choice context characteristics as follows:

$$\boldsymbol{q}_{m,q} = F(\boldsymbol{a} + \boldsymbol{g}' X_q) \tag{19}$$

where $\boldsymbol{q}_{m,q}$ is the structural parameter for nest *m* case *q*,
X_q is a vector of individual and trip related characteristics for case *q*,
$\boldsymbol{a}, \boldsymbol{g}'$ are parameters to be estimated, and
F is a transformation function

Since utility maximization requires that $\boldsymbol{q}_{m,q}$ lie between zero and one, *F* can be any continuous function that maps from a real line to the 0-1 interval. If $\boldsymbol{g}' = 0$ in equation (19), covariance heterogeneity is absent and the COVNL reduces to the NL model. The parameterization of the structural parameter provides additional behavioral appeal to the NL model. The COVNL model retains a simple form and provides closed form expressions for choice probabilities. In its only empirical application to date, the COVNL model was statistically superior to the NL and MNL models, suggesting the potential value of accommodating

covariance heterogeneity across cases in models that allow for correlation among alternatives.

FUTURE DEVELOPMENTS IN CLOSED FORM CHOICE MODELS

Considerable progress has been made in relaxing the independence across alternatives and the homogeneity of error variance across cases within the context of closed form extensions of the multinomial logit model. This progress has dramatically increased the potential to represent complex choice situations using closed form models. Additional developments are likely along three related dimensions.

First, an important aspect of some of these models; specifically, the nested logit, cross-nested logit and generalized nested logit models; is the flexibility of their nesting structure. This flexibility provides more realism in the representation of substitution relationships among alternatives but can result in an extended search and evaluation process. The development of techniques to search intelligently among large numbers of alternatives and provide useful guidance to the analyst would increase the usability of these models in choice contexts with more than a few alternatives.

Second, it is possible to over-parameterize the structural elements of the more complex models causing identification problems. This problem can be addressed, to some extent, by the use of constrained maximum likelihood but may under some circumstances result in non-unique solutions or interpretation difficulties. Additional analysis is required to evaluate the impact of these issues on the usefulness of these models.

Finally, there is opportunity to combine isolated extensions of these closed form models in an integrated framework to yield behaviorally richer, yet computationally tractable models.

Table 1: Cross-Elasticities of Selected GEV models: The Elasticity of Alternative j in response to a Change in Attribute k of Alternative i, X_{ik}

Model	Cross-Elasticity
Multinomial Logit (MNL)	$P_i \boldsymbol{b}_k X_{ik}$
Nested Logit (NL)	$-P_i \boldsymbol{b}_k X_{ik}$ $-\left[P_i + \left(\dfrac{1-\boldsymbol{q}_m}{\boldsymbol{q}_m} \right) P_{i/m} \right] \boldsymbol{b}_k X_{ik}$
Paired Combinatorial Logit (PCL)	$-\left\{ P_i + \left(\dfrac{1-\boldsymbol{q}_{ij}}{\boldsymbol{q}_{ij}} \right) \dfrac{\left[P(ij) \right]\left[P(i|ij)P(j|ij) \right]}{P_j} \right\} \boldsymbol{b}_k X_{ik}$
Cross Nested Logit (CNL)	$-\left\{ P_i + \left(\dfrac{1-\boldsymbol{q}}{\boldsymbol{q}} \right) \sum_m \dfrac{\left[P(m) \right]\left[P(i|m)P(j|m) \right]}{P_j} \right\} \boldsymbol{b}_k X_{ik}$
Generalized Nested Logit (GNL)	$-\left\{ P_i + \sum_m \left(\dfrac{1-\boldsymbol{q}_m}{\boldsymbol{q}_m} \right) \dfrac{\left[P(m) \right]\left[P(i|m)(P(j|m)) \right]}{P_j} \right\} \boldsymbol{b}_k X_{ik}$

REFERENCES

Ben-Akiva, M. (1974). Note on the specification of a Logit Model with Utility Functions that Include Attributes of Competing Alternatives, *Working Paper*, Department of Civil Engineering, MIT, Cambridge, Mass.

Bhat C.R. (1997). A Nested Logit Model with Covariance Heterogeneity, *Transportation Research*, *Part B*, Vol. 31, pp. 11-21.

Bhat, C.R. (1998). Accommodating Flexible Substitution Patterns in Multi-Dimensional Choice Modeling: Formulation and Application to Travel Mode and Departure Time Choice, *Transportation Research*, *Part B*, Vol. 32, No.7, pp. 425-440.

Börch-Supan, A. and V.A. Hajivassiliou (1993). Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models. *Journal of Econometrics*, Vol. 58, No. 3, pp.347-368.

Bresnahan, T.E., S. Stern, and M. Trajtenberg (1997). Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the late 1980s, *Rand Journal of Economics*, Vol. 28, Special Issue, pp. S17-S44.

Cascetta, E., 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, *Transportation and Traffic Theory*, pp. 697-711.

Chu, C. (1989) A Paired Combinatorial Logit Model for Travel Demand Analysis, *Proceedings of the Fifth World Conference on Transportation Research*, Vol. 4, Ventura, CA, 1989, pp. 295-309.

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

Daganzo, C.F. and M. Kusnic, (1993). Two Properties of Nested Logit Model, *Transportation Science*, Vol. 27, pp. 395-400.

Daly, A. and S. Zachary (1979). Improved Multiple Choice Models, in *Determinants of Travel Choice*, Hensher, D.A. and M. Q. Dalvi (eds.), pp. 335-357, Prager, New York.

Gaudry, M. and M. Degenais (1978) The Dogit Model, *Transportation Research*, *Part B*, Vol. 13, No. 2, pp. 105-111.

Hajivassiliou, V. H., and D. McFadden (1990). The Method of Simulated Scores for the Estimation of LDV Models with an Application to External Debt Crises. *Cowles Foundation Discussion Paper 967*, Yale

University.

Johnson, N.L. and S. Kotz (1970). *Distributions in Statistics: Continous Multivariate Distributions*. Chapter 21, John Wiley, New York.

Keane, M. A (1994). Computational Practical Simulation Estimator for Panel Data. *Econometrica,* Vol.62, No.1, pp.95-116.

Koppelman, F.S. and C-H. Wen (2000) The Paired Combinatorial Logit Model: Properties, Estimation and Application, Accepted for Publication, *Transportation Research, Part B*.

McFadden, D. (1973). Conditional Logit Analysis of Quantitative Choice Behavior, in Zaremmbka P. (ed.), *Frontier of Econometrics*, Academic Press, New York.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Choice Response Models without Numerical Integration. *Econometrica,* Vol.57, No.5, pp.995-1026.

McFadden, D. (1975). On Independence, Structure, and Simultaneity in Transportation Demand Analysis, Working Paper No. 7511, Urban Travel Demand Forecasting Project, Institute of Transportation and Traffic Engineering, University of California, Berkeley.

McFadden D. (1978). Modeling the Choice of Residential Location, *Transportation Research Record*, No. 672, pp. 72-77.

Small, K. (1987). A Discrete Choice Model for Ordered Alternatives, *Econometrica*, Vol. 55, pp. 409-424.

Swait, J. and M. Ben-Akiva (1987). Empirical Test of a Constrained Choice Discrete Model : Mode Choice in Sao Paulo, Brazil *Transportation Research,* Vol. 21B, No. 2, 103-116.

Swait, J. and W. Adamowicz (1996). The Effect of Choice Environment and Task Demands on Consumer Behavior: Discriminating between Contribution and Confusion, Working Paper, Department of Rural Economy, University of Alberta.

Swait, J. (2000). Choice Set Generation within the Generalized Extreme Value Family of Discrete Choice Models, Working Paper, University of Florida, Gainesville, FL.

Voshva, P. (1998). Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area, *Transportation Research Record*, No.1607, pp. 6-15.

Voshva, P. (1999). E-Mail to the lead author describing the FNL model, October.

Wen, C.H. and F. S. Koppelman (1998). Paired Combinatorial Logit: Properties, Estimation and Application, Presented at the Annual Conference of Transportation Research Board, Washington D.C., January 1998.

Williams, H.C.W.L. (1977). On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit, *Environment and Planning*, 9A, 285-344.

Williams, H.C.W.L. and J. de D. Ortúzar (1982). Behavioural theories of dispersion and the mis-specification of

travel demand models. *Transportation Research 16B*, 167-219.

Wen, C.H. and Koppelman, F.S., (2000). A Generalized Nested Logit Model, Presented at the Annual Conference of The Transportation Research Board, Washington D.C., January..