Recent methodological developments in discrete choice models

M. Bierlaire*

April 15, 2007

Report TRANSP-OR 070415
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

*École Polytechnique Fédérale de Lausanne, Transport and Mobility Laboratory, Station 18, CH-1015 Lausanne, Switzerland. E-mail: michel.bierlaire@epfl.ch

Abstract

This short document contains lecture notes of an invited plenary lecture made during the conference "Metodi, modelli e tecnologie dell'informazione a supporto delle decisioni", organized by Università di Napoli Federico II on September 30, 2006, on the island of Procida, Italy.

1 Introduction

The role of discrete choice models in decision-making support has significantly grown during the past 20 years. This is particularly true in the context of marketing and transportation, where it is critical to understand and forecast choice behavior in a detailed way. In this short document, we provide an overview of recent methodological developments in this area. The multinomial logit (MNL) model is probably the most popular random utility model, due to its relative simplicity. However, its derivation is based on strong independence assumptions. Namely, error terms in the utility functions are supposed to be independent across alternatives and individuals. Such assumptions are not valid in many concrete contexts. The family of Multivariate (or Generalized) Extreme Value (MEV) models relax the assumption of independence across alternatives. This family, proposed by McFadden (1978) includes the nested logit model, the cross nested logit model and the network MEV model. Convenient because of the closed form of the probability formula, MEV models suffer from some limitations, one of them being their intrinsic homoscedasticity. Mixtures of MEV models can be derived to overcome these limitations (McFadden and Train, 2000). While mixing provides a great deal of flexibility, it significantly complicates the estimation of the model, as the probability model has no closed form any more. Therefore, simulation is required, which is computationnally intensive. We also touch upon the issue of testing if the mixing distribution is adequate (Fosgerau and Bierlaire, 2007).

2 Random utility models

Random utility models are derived from the concept of utility maximization. Decision-makers are assumed to be rational, and to perform a choice in order to maximize a quantity, called *utility*, associated with each of the alternatives under consideration. The utility is modeled by a random variable, in order to account for the many sources of uncertainty in the decision process itself, and in the methodological assumptions. Discrete choice models are based on the assumption that the set of alternatives considered by the decision-maker, or *choice set*, is finite and discrete. If the choice set of decision-maker n is denoted by $\mathcal{C}_n$, a discrete choice model provides the probability that the decision-maker chooses an alternative i within $\mathcal{C}_n$, given the vector of explanatory variables x_n, combining the socio-economic characteristics of decision-maker n, and the attributes of each alternative for this individual, that is

$$P_n(i|x_n, \mathcal{C}_n). \tag{1}$$

Random utility models are discrete choice models such that

$$P_n(i|x_n, \mathcal{C}_n) = \Pr(U_{in}(x_n) \geq U_{jn}(x_n), \forall j \in \mathcal{C}_n), \tag{2}$$

where $U_{in}(x_n)$ is the random variable representing the utility associated by decision-maker n with alternative i. Writing

$$U_{in}(x_n) = V_{in}(x_n) + \varepsilon_{in}, \tag{3}$$

where $V_{in}(x_n)$ is a real, deterministic number, capturing the *systematic part of the utility*, and ε_{in} is a random variable called the *error term*, specific models can be derived based on explicit assumptions about $V_{in}(x_n)$ and ε_{in}.

The Multinomial Logit (MNL) model, probably the most popular random utility model, is based on the assumption that the error terms ε_{in} are independently, identically and extreme value (EV) distributed. In this case, (2) becomes

$$P_n(i|x_n, \mathcal{C}_n) = \frac{e^{V_{in}(x_n)}}{\sum_{j \in \mathcal{C}_n} e^{V_{jn}(x_n)}}. \tag{4}$$

We refer the reader to Ben-Akiva and Lerman (1985) and Train (2003) for a detailed discussion on discrete choice models.

3 Complex error structures

In order to relax the independence assumption associated with the MNL model, we consider the utilities as a vector of J random variables, where J is the number of alternatives in the choice set $\mathcal{C}$. For the sake of simplification, we assume that $\mathcal{C}$ is identical for all decision-makers, and we drop subscript n. We obtain

$$\begin{pmatrix} U_1(x) \\ \vdots \\ U_J(x) \end{pmatrix} = \begin{pmatrix} V_1(x) \\ \vdots \\ V_J(x) \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_J \end{pmatrix} \tag{5}$$

that is, $U(x) = V(x) + \varepsilon$. The *probit* model is based on the assumption that ε has a multi-variate normal distribution. However, for models with more than 4 or 5 alternatives, the use of such a distribution involves the evaluation of a multifold integral which is computationally infeasible. Therefore, multivariate extreme value (MEV) distributions and mixtures of MEV models are usually preferred.

3.1 Multivariate extreme value distribution

The Multivariate Extreme Value model was first proposed by McFadden (1978), under the name "Generalized Extreme Value" model. In order to avoid any confusion with the univariate Generalized Extreme Value distribution used in the statistics literature (see Kotz and Nadarajah, 2001), we prefer to refer to this model as *Multivariate Extreme Value* (MEV).

Each instance of the MEV family is derived from a μ-MEV function. It is a differentiable function

$$G : \mathbb{R}_+^{J_n} \to \mathbb{R}_+, \tag{6}$$

where J_n is the number of alternatives in the choice set $\mathcal{C}_n$. The function G is used to define a CDF and, consequently, a choice model. The CDF of a MEV distribution is given by

$$F(\varepsilon_1, \ldots, \varepsilon_{J_n}) = e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_{J_n}})}. \tag{7}$$

In order for F to be a CDF, the μ-MEV function G must have the following properties:

1. $G(y) > 0$ for all $y \in \mathbb{R}_+^{J_n}$,
2. G is homogeneous of degree $\mu > 0$, that is $G(\lambda y) = \lambda^\mu G(y)$, for $\lambda > 0$,
3. $\lim_{y_i \to +\infty} G(y_1, \ldots, y_i, \ldots, y_{J_n}) = +\infty$, for each $i = 1, \ldots, J_n$,
4. the mixed partial derivatives of G exist and are continuous. Moreover, the kth partial derivative with respect to k distinct y_i is non-negative if k is odd and non-positive if k is even that is, for any distinct indices $i_1, \ldots, i_k \in \{1, \ldots, J_n\}$, we have

$$(-1)^k \frac{\partial^k G}{\partial y_{i_1} \ldots \partial y_{i_k}}(y) \le 0, \ \forall y \in \mathbb{R}_+^J. \tag{8}$$

We say that a function satisfying this property is *MEV-differentiable.*

Note that the original formulation by McFadden (1978) is based on $\mu = 1$. The generalisation with $\mu > 0$ has been derived by Ben-Akiva and François (1983).

The marginal distribution is obtained by sending all ε_j but one to infinity, that is $\varepsilon_j \to +\infty$, $\forall j \neq i$. We have

$$\begin{aligned} F(+\infty, \ldots, e^{-\varepsilon_i}, \ldots, +\infty) &= e^{-G(0,\ldots,e^{-\varepsilon_i},\ldots,0)} \\ &= e^{-e^{-\mu\varepsilon_i} G(0,\ldots,1,\ldots,0)} \\ &= e^{-\alpha e^{-\mu\varepsilon_i}} \\ &= e^{-e^{-\mu\varepsilon_i + \ln\alpha}}, \end{aligned} \tag{9}$$

where $\alpha = G(0, \ldots, 1, \ldots, 0)$. This is the CDF of an extreme value distribution, with location parameter $\eta = -\ln\alpha/\mu$ and scale parameter μ, justifying the fact that F is the CDF of a multivariate extreme value distribution.

The MEV probability model can be derived from (5) and (7), to obtain

$$P_n(i) = \frac{e^{V_{in}} G_i(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})}{\mu G(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})}, \tag{10}$$

where G_i denotes the partial derivative of G with respect to the ith variable. As G is homogeneous, Euler's theorem may be invoked to obtain an equivalent formulation of the MEV probability model:

$$P_n(i) = \frac{e^{V_{in} + \ln G_i(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})}}{\sum_j e^{V_{jn} + \ln G_j(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})}}. \tag{11}$$

Formulation (11) is interesting because it has a similar structure as the MNL model. Indeed, it can be interpreted as a MNL model, where each systematic utility V_i is shifted by $\ln G_i(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})$.

The expected maximum utility is given by

$$V_{\mathcal{C}} = \frac{\ln G(e^{V_{1n}}, \dots, e^{V_{Jn}}) + \gamma}{\mu}, \tag{12}$$

where γ is Euler's constant (≈ 0.5772). Differentiating this expression with respect to V_{in}, we obtain that

$$\frac{\partial V_{\mathcal{C}}}{\partial V_i} = \frac{e^{V_{in}} G_i(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})}{\mu G(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{Jn}})} = P_n(i). \tag{13}$$

The derivation of the Multinomial Logit model, the Nested Logit model and the Cross-Nested model from the MEV class can be found in various sources, including Ben-Akiva and Lerman (1985), Wen and Koppelman (2001), Ben-Akiva and Bierlaire (2003) and Bierlaire (2006).

The main advantages of MEV models is the closed form of the probability model (contrarily to the probit model), and the great deal of flexibility in the correlation structure. However, there is no free lunch, and various issues must be addressed. We cite a few that have recently been studied.

In many applications, the correlation structure of the model is known. In a probit model, this can be reflected in the variance-covariance matrix. In a MEV model, the formulation is based on the μ-MEV function and the variance-covariance matrix does not appear explicitly. Inspired by the work of Papola (2004) and Papola and Marzano (2005), Abbe et al. (2007) have described how to derive a cross-nested logit model from a given correlation structure.

Also, the derivation of new instances of the MEV family is not straightforward and requires heavy proofs. Daly and Bierlaire (2006) have provided a series of theorem for MEV calculus, allowing to recursively derive new MEV models from existing ones. They propose a network representation that simplifies the work on the analyst to represent a desired correlation structure, similar to the trees used for nested logit models. A MEV model defined based on such a network structure is called a *network MEV* model.

In the presence of choice-based sampling strategies for data collection, the multinomial logit model has the convenient property that consistent estimates of all parameters but the constants can be obtained from an Exogenous Sample Maximum Likelihood (ESML) estimation (result by McFadden, reported by Manski and Lerman, 1977). Unfortunately, it does not hold in general for MEV models. Bierlaire et al. (2006) analyze how the estimation should be performed in practice in order to keep it as simple as possible.

3.2 Mixtures of MEV

In statistics, a *mixture density* is a pdf which is a convex linear combinations of other pdf's. If $f(\varepsilon, \theta)$ is a pdf, and if $w(\theta)$ is a nonnegative function such that $\int_\theta w(\theta) d\theta = 1$ then

$$g(\varepsilon) = \int_\theta w(\theta) f(\varepsilon, \theta) d\theta$$

is also a pdf. We say that g is a mixture of f. In particular, if f is the pdf of a MEV model, it is a mixture of MEV models.

Discrete mixtures are also possible. If $f(\varepsilon, \theta)$ is a pdf, and if w_i, $i = 1, \ldots, p$ are nonnegative weights such that $\sum_{i=1}^{p} w_i = 1$ then

$$g(\varepsilon) = \sum_{i=1}^{p} w_i f(\varepsilon, \theta_i)$$

is also a pdf. We say that g is a discrete mixture of f. Again, if f is the pdf of a MEV model, it is a discrete mixture of MEV models.

Error component models are typical example of mixtures of MNL models. In this case, the correlation between two (or more) alternatives is explicitly captured by a common error component. In a trinomial case, for instance, we may capture the correlation between alternatives 1 and 2 as follows:

$$\begin{array}{rclll} U_1 & = & V_1 & +\xi & +\varepsilon_1 \\ U_2 & = & V_2 & +\xi & +\varepsilon_2 \\ U_3 & = & V_3 & & +\varepsilon_3, \end{array} \tag{14}$$

where ε_i, $i = 1, 2, 3$, are i.i.d. EV distributed, and ξ is a random variable with a given pdf $f(\cdot)$. Therefore, if ξ was given, the probability model would be a MNL, that is

$$P(1|\xi) = \frac{e^{V_1+\xi}}{e^{V_1+\xi} + e^{V_2+\xi} + e^{V_3}}. \tag{15}$$

As ξ is not given but distributed, it has to be integrated out, to obtain

$$P(1) = \int_\xi \frac{e^{V_1+\xi}}{e^{V_1+\xi} + e^{V_2+\xi} + e^{V_3}} f(\xi) d\xi. \tag{16}$$

In general, equation (16) has no closed form, and simulation techniques are required (see Train, 2003).

This error component specification provides a great deal of flexibility. Actually, McFadden and Train (2000) have shown that *under mild regularity conditions, any discrete choice model derived from random utility maximization has choice probabilities that can be approximated as closely as one pleases by a Mixed MNL model.*

A direct extension is the *factor analytic* specification, proposed by Ben-Akiva and Bolduc (1996). Actually, they have developed for logit models an idea by McFadden (1984) to reduce the dimensionality of probit models (see also Bierlaire, 2005a).

Another typical use of the mixture of models is to capture unobserved heterogeneity, with the use of random coefficients in the model specification.

We refer the reader to McFadden and Train (2000), Walker (2001), Hensher and Greene (2003), Ben-Akiva and Bierlaire (2003) and Train (2003) for more details. Mixtures of MEV models have been used by Bhat and Guo (2004) and Hess et al. (2005a).

4 Testing

In the context of mixture models, an important issue is the choice of a specific distribution for the random parameters. Actually, various pieces of research have demonstrated that an inappropriate choice of the distribution may lead to serious biases in model forecast and in the estimated mean of random parameters. A noticeable example is the Normal distribution, used as a default for many applications. Hess et al. (2005b) discuss wrong interpretations of willingness-to-pay indicators when normal distributions are considered. Fosgerau (2006) looks at various distributions and concludes that a bad choice may lead to extreme biases. Hess and Axhausen (2005) have examined how well a wide range of parametric distributions can reproduce given target distributions, which are constructed to reflect common assumptions about taste variation in transport demand models.

Fosgerau and Bierlaire (2007) propose a practical test, based on seminonparametric techniques. The term seminonparametric distinguishes a certain class of models from parametric, nonparametric and semiparametric models. Semiparametric models are a hybrid between parametric and nonparametric models. Seminonparametric models use series approximations to approximate functions such as densities.

The test checks if a random parameter ω of a discrete choice model indeed follows a base distribution with CDF F and density f. The true distribution may be different from F. Denoting the true CDF by G and its density by g, the distribution G can be rewritten in terms of F as

$$G(\omega) = Q(F(\omega)),$$

where Q is a monotone function from $[0, 1]$ to $[0, 1]$. As such, Q is a CDF for a stochastic variable on the unit interval. Differentiating, we obtain

$$g(\omega) = q(F(\omega))f(\omega).$$

Now, q can be approximated in a seminonparametric fashion as

$$q(x) \approx \frac{1}{K} q_N^2(x), \tag{17}$$

where

$$q_N(x) = 1 + \sum_{k=1}^{N} \delta_k L_k(x), \tag{18}$$

$$K = \int_{-\infty}^{+\infty} q_N^2(F(\omega)) f(\omega) d\omega \tag{19}$$

is a normalizing constant, and L_k is the transformed Legendre polynomials proposed by Bierens (2005) (see Fosgerau and Bierlaire, 2007 for details).

If β is a parameter of a discrete choice model, the probability for alternative i to be chosen in choice set $\mathcal{C}_n$ is given by

$$P_n(i|\mathcal{C}_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, \mathcal{C}_n) g(\beta) d\beta,$$

where $P_n(i|\beta, \mathcal{C}_n)$ is a closed form model, such as the MEV model. Then,

$$\begin{aligned} P_n(i|\mathcal{C}_n) &\approx \frac{1}{K} \int_{-\infty}^{+\infty} P_n(i|\beta, \mathcal{C}_n) q_N^2(F(\beta)) f(\beta) d\beta \\ &= \frac{1}{K} \int_{0}^{1} P_n(i|F^{-1}(z), \mathcal{C}_n) q_N^2(z) dz, \end{aligned}$$

using the change of variables $z = F(\beta)$. This integral is approximated by Monte-Carlo simulation, and the term $F^{-1}(z)$ corresponds to the draws of the base distribution.

Under the null hypothesis that the base distribution is the true distribution, we have $f = g$, which implies that q is identically 1 and thus that $\delta_k = 0$, for all k in (18). Then the model

$$P_n(i|\mathcal{C}_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, \mathcal{C}_n) g(\beta) d\beta, \tag{20}$$

is equivalent to the model

$$P_n(i|\mathcal{C}_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, \mathcal{C}_n) f(\beta) d\beta. \tag{21}$$

Because the two models are nested, a standard likelihood ratio test for nested hypotheses is appropriate to test the null hypothesis that $f = g$.

The concept is illustrated on synthetic and real data in the paper by Fosgerau and Bierlaire (2007). The test is implemented in the software package biogeme, briefly described below.

5 Biogeme

Biogeme (Bierlaire, 2003,Bierlaire, 2005b) is an open source software package designed for the maximum likelihood estimation of a wide variety of MEV models and mixtures of MEV models, using various nonlinear optimization algorithms.

All information relative to BIOGEME is maintained at

`biogeme.epfl.ch`.

The archives of the users group are at

`groups.yahoo.com/group/biogeme`

The current version of Biogeme has the following main features:

MEV models The following models can be estimated:

- multinomial logit models,
- nested logit models,
- cross-nested logit models,
- network-MEV models.

Binary probit models can be estimated.

Continuous and discrete mixtures of MEV models Normally distributed and uniformly distributed random parameters can appear in the utility functions, as well as random parameters with discrete distributions. In the presence of parameters with continuous distributions, Simulated Maximum Likelihood is performed, using either pseudo-random numbers, quasi-random numbers generated from Halton sequences, or the Modified Latin Hypercube Sampling strategy.

Panel data It is possible to estimate individual-specific random parameters.

No scaling issue An important feature of Biogeme is that the specification of the utility function is completely independent from the distribution of the error term. It has the advantage that the user does not need to worry about scaling problems when constraining parameters appearing in different nests of a nested or cross-nested logit model, for instance. So the issues raised by Koppelman and Wen (1998a) and Koppelman and Wen (1998b) are not relevant in this context.

Nonlinear utility functions The utility functions do not need to be linear-in-parameters, like in most standard packages. A wide range of nonlinear specifications are allowed. No derivative is requested from the user. As a byproduct, log-normally distributed parameters can also be specified.

We encourage everybody interested in discrete choice models to use this free software.

References

Abbe, E., Bierlaire, M. and Toledo, T. (2007). Normalization and correlation of cross-nested logit models, *Transportation Research Part B: Methodological* . Forthcoming, accepted for publication.

Ben-Akiva, M. and Bierlaire, M. (2003). Discrete choice models with applications to departure time and route choice, *in* R. Hall (ed.), *Handbook of Transportation Science, Second Edition*, Kluwer Academic Publishers, Boston/Dordrecht/London, pp. 7–37.

Ben-Akiva, M. and Bolduc, D. (1996). Multinomial probit with a logit kernel and a general parametric specification of the covariance structure. Paper presented at the 3rd Invitational Choice Symposium, Columbia University.

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

Ben-Akiva, M. and François, B. (1983). μ homogeneous generalized extreme value model, *Working paper*, Department of Civil Engineering, MIT, Cambridge, Ma.

Bhat, C. R. and Guo, J. Y. (2004). A mixed spatially correlated logit model: formulation and application to residential choice modeling, *Transportation Research Part B: Methodological* **38**(2): 147–168.

Bierens, H. J. (2005). Semi-nonparametric modeling of densities on the unit interval, with applications to censored mixed proportional hazard models and ordered probability models: Identification and consistency results, *Econometric Society World Congress*.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland. www.strc.ch.

Bierlaire, M. (2005a). Capturing complex transportation behaviors with advanced random utility models, *in* H. van Zuylen (ed.), *The reliability of traveling and the robustness of transport systems*, Delft, The Netherlands, TRAIL Research School.

Bierlaire, M. (2005b). An introduction to BIOGEME version 1.4. biogeme.epfl.ch.

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model, *Annals of operations research* **144**(1): 287–300.

Bierlaire, M., Bolduc, D. and McFadden, D. (2006). The estimation of generalized extreme value models from choice-based samples, *Technical Report TRANSP-OR 060810*, Transport and Mobility Laboratory, ENAC, EPFL, Lausanne, Switzerland. Download from transp-or.epfl.ch.

Daly, A. and Bierlaire, M. (2006). A general and operational representation of generalised extreme value models, *Transportation Research Part B: Methodological* **40**(4): 285–305.

Fosgerau, M. (2006). Investigating the distribution of the value of travel time savings, *Transportation Research Part B: Methodological* **40**(8): 688–707.

Fosgerau, M. and Bierlaire, M. (2007). A practical test for the choice of mixing distribution in discrete choice models, *Transportation Research Part B: Methodological* . Forthcoming, accepted for publication.

Hensher, D. and Greene, W. (2003). The mixed logit model: the state of practice, *Transportation* **30**(2): 133–176.

Hess, S. and Axhausen, K. W. (2005). Distributional assumptions in the representation of random taste heterogeneity, *Proceedings of the 5th Swiss Transportation Research Conference*. www.strc.ch.

Hess, S., Bierlaire, M. and Polak, J. (2005a). Capturing correlation and taste heterogeneity with mixed GEV models, *in* R. Scarpa and A. Alberini (eds), *Application of simulation methods in environmental and resource economics*, Vol. 6 of *The Economics of Non-Market Goods and Resources*, Springer.

Hess, S., Bierlaire, M. and Polak, J. (2005b). Estimation of value of travel-time savings using mixed logit models, *Transportation Research Part A* **39**(3): 221–236.

Koppelman, F. S. and Wen, C.-H. (1998a). Alternative nested logit models: Structure, properties and estimation, *Transportation Research Part B: Methodological* **32**(5): 289–298.

Koppelman, F. S. and Wen, C. H. (1998b). Nested logit models: which are you using?, *Transportation Research Record* **1645**: 1–7.

Kotz, S. and Nadarajah, S. (2001). *Extreme Value Distributions: Theory and Applications*, Imperial College Press.

Manski, C. and Lerman, S. (1977). The estimation of choice probabilities from choice-based samples, *Econometrica* **45**: 1977–1988.

McFadden, D. (1978). Modelling the choice of residential location, *in* A. Karlquist *et al.* (ed.), *Spatial interaction theory and residential location*, North-Holland, Amsterdam, pp. 75–96.

McFadden, D. (1984). Econometric analysis of qualitative response models, *in* Z. Friliches and M. Intriligator (eds), *Handbook of Econometrics II*, Elsevier Science Publishers.

McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response, *Journal of Applied Econometrics* **15**(5): 447–470.

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

Papola, A. and Marzano, V. (2005). A specification procedure of the CNL model reproducing any homoskedastic covariance matrix, *Proceedings of the European Transport Conference*, Strasbourg, France.

Train, K. (2003). *Discrete Choice Methods with Simulation*, Cambridge University Press. http://emlab.berkeley.edu/books/choice.html.

Walker, J. L. (2001). *Extended discrete choice models: integrated framework, flexible error structures, and latent variables*, PhD thesis, Massachusetts Institute of Technology.

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