



Logit-based approaches for modelling dominated choice alternatives

Francisco Martinez

Department of Civil Engineering at the University of Chile fmartine@ing.uchile.cl

Ennio Cascetta and Francesca Pagliara

*Department of Transportation Engineering – University of Naples Federico II
cascetta@unina.it; fpagliar@unina.it*

Michel Bierlaire

*Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering
EPFL Ecole Polytechnique Fédérale de Lausanne
michel.bierlaire@epfl.ch*

Kay W. Axhausen

*IVT – ETH
axhausen@ivt.baug.ethz.ch*

Abstract: *The objective of this paper is to specify and estimate a Constrained Multinomial Logit model with dominance variables. Estimation results will be compared with a simple Multinomial Logit model including dominance variables as well.*

Keywords: *random utility theory, logit models, dominance variables, cut-offs.*

1. Introduction

Random utility models are widely used to analyze choice behaviour and predict choices among discrete alternatives in a given set. These models are based on the assumption that an individual's preference for the available alternatives can be described with a utility function and that the individual selects the alternative with the highest utility (Ben-Akiva and Lerman, 1985). The traditional formulation of logit models applied to transport demand assumes compensatory (indirect) utilities based on the trade-off between attributes. Some authors have criticized this approach because it fails to recognize attribute thresholds in consumers' behavior, as for example the process of elimination-by-aspects (EBA), or a more generic feasible choices domain where such compensatory strategy is contained.

A different strategy has been proposed by Cascetta and Papola (2001) and Martinez et al. (2005), which makes those choice alternatives out of the feasible domain available but undesirable. This approach has the advantage that the model is applied to the entire set of choices, thus gaining on efficiency by avoiding the explicit identification of choice sets for every individual, and secondly, obtaining a model with better properties for the calculation of equilibrium or optimum conditions (see Martinez et al., 2006). Based on this approach, in this paper the Constrained Multinomial Logit (CMNL) model will be specified, combining the multinomial logit model with a binomial logit factor that represents soft cut-offs.

2. The methodology

Under the principle of rationality, some alternatives may not be considered because they are dominated by another alternative.

Dominance effects can be implemented by variables in the utility function, defined by combining the rules defined by Cascetta and Papola (2005) and Cascetta et al. (2007), which generate dominance values that will be assigned to each alternative.



In this paper a different method is introduced using cut-off factors, instead of assigning dominance values to alternatives. These cut-off factors will represent the probability of an alternative for being dominated by other alternatives.

The proposed approach is applied using the revealed preference survey conducted in 2005 in the canton of Zurich in Switzerland, covering the mobility and moving biography of the respondents.

In Cascetta et. al. (2007) the residential location model was specified as a Multinomial Logit (MNL) model using the following linear utility function:

$$U_{d \in I}^* = U_d + \varepsilon_d = \sum_{n \in N} \alpha_n X_{dn} + \sum_{k \in K} \beta_k Y_{dk} + \varepsilon_d \quad (1)$$

where X_{dn} are the values of the compensatory utility variables in zone d and Y_{dk} are the dominance variables on zone d ; α_n and β_k are the respective parameters; I , N and K are, respectively, the sets of zones (residential location options), utility variables and dominance variables. The last one is the random term assumed distributed identically and independently distributed Gumbel. Then the location choice multinomial logit model is:

$$p(d) = \frac{\exp(U_d)}{\sum_{j \in I} \exp(U_j)} \quad (2)$$

Applying the cut-off method proposed by Martinez et al. (2005) and the dominance variables proposed by Cascetta et al. (2007), the CMNL model is specified.

The following cut-off factor is defined:

$$\phi(Y_{dk}) = \frac{1}{1 + \exp\{\omega(Y_{dk} - Y_k + \rho)\}} \quad (3)$$

where Y_k is the cut-off level, or the level of dominance above which location choices become irrelevant so they are detracted or ignored from the choice set, or equivalently, their individuals utility is so low or negative that these locations are not considered by the individual. The cut-off function (4) can be interpreted as a binomial logit model where the alternatives options are if those alternatives that violate the maximum dominance level are included or not in the choice set.

Then the utility function is defined as:

$$U_d = V_d + \Phi_d = \sum_{l \in L} \alpha_l X_{dl} + \frac{1}{\mu} \sum_{k \in K} \ln \phi(Y_{dk}) \quad (4)$$

If any of the K criterion for dominance takes a value $Y_{dk} \geq Y_k$, then the cut-off tends to zero and the utility falls to minus infinity, thus making the location option irrelevant although still feasible.

The CMNL model will be calibrated (Bierlaire, 2007) by estimating the set of parameters that best fits the available sample of the observed residential choices. The calibration procedure starts defining the level of disaggregation of the model and parameter estimates. The most disaggregate level considers the estimation of the following parameter's vector $(\alpha_{nl}, \omega_{nk}, \rho_{nk})_{n \in N, l \in L, k \in K}$, which includes specific parameters for each individuals socioeconomic category n , compensating utility variables l , and for each variable cut-off k . This definition is highly dependent on the available data.

Bibliography

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



- Bierlaire, M. (2007), An introduction to BIOGEME (Version 1.4), EPF Lausanne, Lausanne 2007)
<http://transp-or2.epfl.ch/biogeme/doc/tutorial.pdf>.
- Cascetta E., Pagliara F., Axhausen K. W.(2007), The use of dominance variables in spatial choice modelling. *Proceedings of the 11th World Conference on Transport Research*, University of California at Berkely, June.
- Cascetta, E., Papola, A. (2001), Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand, *Transportation Research C* 9: 249-263.
- Cascetta, E., Papola, A. (2005), Dominance among alternatives in random utility models: a general framework and an application to destination choice. *Proceedings of the European Transport Conference*, Strasbourg, October.
- Martínez F., Aguila F., Hurtubia R (2005), The Constrained Multinomial Logit Model: A Semi-Compensatory Choice Model. *Proceedings of the 8th International Conference on Computers in Urban Planning and Urban Management* London.