Advantages of latent class choice models over continuous Mixed Logit models

Stephane Hess, Institute for Transport Studies - University of Leeds
Moshe Ben-Akiva, Massachusetts Institute of Technology
Dinesh Gopinath, ChoiceStream
Joan Walker, University of California at Berkeley

Abstract
Latent class model structures present a flexible extension of standard choice modelling approaches in the context of the representation of taste heterogeneity. A latent class model divides the population into a number of classes with differences in sensitivities across classes. With the class membership being treated as a latent component, a probabilistic class-allocation model is used, meaning that an individual has a non-zero probability of belonging to each of the different classes. The class-membership probabilities are a function of socio-demographic attributes of the decision-maker, meaning that any taste heterogeneity across respondents can be linked directly to variations in socio-demographic attributes across individuals. This can be a major asset in the interpretation of model results as well as in forecasting. While to some extent taste heterogeneity can also be linked to socio-demographic indicators in a more standard modelling framework, the degree of flexibility (in terms of the extent of heterogeneity) is generally quite limited, and the specification search can be a very tedious task.

Despite these very appealing characteristics, latent class approaches are used relatively rarely in the field of travel behaviour research, and have been largely upstaged by the increasing popularity of the Mixed Logit model. However, it should be said that latent class models have very significant advantages in interpretation over the Mixed Logit model. Indeed, being able to link taste heterogeneity to socio-demographic indicators is clearly preferable to simply knowing that a given sensitivity follows a certain (assumed) random distribution in the sample population.

This paper builds on ongoing work by the authors in further developing latent class methodology and highlighting potential advantages over continuous mixture models in terms of flexibility as well as interpretation. We first illustrate once again how the marginal utilities (and by extension willingness-to-pay indicators such as the valuation of travel time savings) are a direct function of socio-demographic attributes used in the class-allocation model and show how latent class models make less restrictive shape assumptions than is the case with most Mixed Logit specifications. As a first contribution, we then derive formulae for the correlation between individual taste coefficients in latent class structures and show how this correlation similarly is a function of the socio-demographic attributes used in the class-allocation model. From this, an analyst can for example conclude that for specific subgroups in the population, the time and cost coefficients are negatively correlated, while for others, the correlation may be positive. This is a crucial advantage over the Mixed Logit model, which only produces a fixed measure of the correlation between two randomly distributed coefficients. As a next step, we show that the same principle applies to the elasticities in latent class models, where these can again be expressed as a function of the socio-demographic attributes used in the class-allocation model.

In the applied part of the paper, we make use of stated choice data for departure time and travel mode collected in the West Midlands region of the United Kingdom in 2003. Our results shows that the latent class model obtains significant gains over the basic MNL model, where these gains in log-likelihood are comparable to those obtained by a Mixed Logit model with a fully specified covariance structure.
between random coefficients. However, the real advantages come at the interpretation stage. Here, our analysis illustrates the relationship between socio-demographic indicators and the covariance structure and elasticities in a latent class model. This information is not only useful for the analysis of taste heterogeneity but can also provide significant advantages in forecasting. Additionally, the results show significant differences in the covariance results between the latent class and Mixed Logit models.