Simultaneously accounting for inter-alternative correlation and taste heterogeneity among long distance travelers using mixed nested logit (MXNL) model so as to improve toll road traffic and revenue forecast

Collins Teye\textsuperscript{a,b}, Peter Davidson\textsuperscript{a,*}, Rob Culley\textsuperscript{a}

\textsuperscript{a}Peter Davidson Consultancy, UK,  
\textsuperscript{b}University of Sydney, Business School, Australia

Abstract

This paper investigates the potential of using Mixed Nested Logit (MXNL) models to simultaneously account for inter-alternative correlation, taste heterogeneity and the distribution of willingness to pay for toll roads, using a Stated Preference dataset from toll route choice experiments conducted during a recent toll route study in Nigeria. By Mixed Nested Logit model we mean the model which combines the mixed logit with the nested logit estimated simultaneously. Results reveal the presence of both correlation (addressed by the nested logit model) and different taste heterogeneity (addressed by the mixed logit model). The estimation results for the combined mixed nested logit model is presented compared with individual estimation results for nested logit on its own, mixed logit on its own and multinomial logit. This paper is unique in that there does not seem to be much work in using this combined mixed nested logit model approach to understanding long distance travellers’ behaviour in the context of road pricing. This paper opens up this area for investigation and shows the additional explanation we can potentially achieve to improve our models forecasting ability.

Keywords: Inter-alternative, correlation, taste-heterogeneity, mixed, nested, logit, toll, distance.

\* Corresponding author. Tel.: +44 -1442891665.  
E-mail address: mail@peter-davidson.com
1. Introduction

The Multinomial logit (MNL) for the past several years has been an effective tool for analysing individual travel behaviour and appraising transport schemes (McFadden, 1974; Hensher, 2001a; Train, 2003; Hess et al, 2005). However, the assumptions underlying this model resulted in the so called IIA (Independent of Irrelevant Alternatives) property. Although this property can be very beneficial in certain applications such as reduction in data and estimation costs, if the IIA reflects reality (Train 2009), it places severe limitations on the ability of the MNL model to produce expected or intuitive results under many applications (Train, 2009; Bierlaire, 2006b; Greene and Hensher, 2003; Bhat, 2003; Hensher and Green).

In Teye, et al, (2013) they classified many of the research work directed at reducing or eliminating the effects of the IIA property of the MNL model into three groups; the first set of models they noted seeks to account for the similarities in the unobserved utilities relating to the alternatives (inter-alternative correlation). These models largely belong to the GEV (generalised extreme value) family of models and include multinomial (McFadden, 1974), Nested and (Daly and Zachary, 1978; Ben-Akiva, and Lerman, 1985) and Cross nested logit (Vovsha, 1997; Papola, 2000 and 2004; Bierlaire, 2001). The second group of models look at the existence of taste heterogeneity among travellers. Models accounting for this phenomenon include the mixed multinomial logit (McFadden and Train, 2000; Bhat 2001, 2003; Hensher and Green, 2001, 2003; Hess et al, 2006; Davidson et al, 2012) and Latent class logit (Green and Hensher, 2003; Green et al, 2006; Hess et al, 2011; Davidson et al, 2012). The third group of models seeks to account simultaneously for both the inter-alternative correlation and taste heterogeneity. Models of this type involve mixed GEV models (Hess, Bierlaire & Polak, 2005) and Latent Class GEV models (Teye et al, 2013).

The weakness in the MNL model was demonstrated in Hess et al, (2005) where they showed that not properly accounting for existence of correlation across choice alternatives due to unobserved attributes could lead to unrealistic substitution patterns, and a misleading forecast of demand. Indeed, the cross elasticity formula for the MNL model (Teye et al, 2013) implies that a change in an attribute of an alternative changes the probabilities for all the other alternatives by the same percentage. That is, there are no common unobserved factors affecting the utilities of the various alternatives, making the MNL unsuitable for several applications. For example, the MNL is unsuitable for predicting mode shares when a decision-maker assigns a higher utility to all public transport (e.g, train, bus) modes because of the opportunity to socialize as opposed to using a car. However, the direct and cross elasticity formulae for the Nested Logit (NL) model (Teye et al, 2013) are functions of the structural parameters. These structural parameters are used to measure the degree of correlation among the alternatives in the same nest (Train, 2009). This allows alternatives within the same nest to become better substitutes of each other than those outside the nest (Teye et al, 2013). Thus with appropriate structuring of the alternatives, the NL has the potential to account for the existence of common unobserved factors affecting the utilities of the various alternatives.

Both the MNL and NL models maintain homogeneity in responsiveness to the attributes of alternatives across individuals. That is, every individual in the sampled population gives the same weight or importance to each of the attributes in the choice process. This assumption is clearly not plausible in reality as you would expect different people to place different emphasis (weights) on the various attributes. For example, a poor person may think that travel cost is very important, but a millionaire probably wouldn’t care much about cost. It therefore seems desirable to allow different individuals to have different weights. One of the ways of dealing with taste heterogeneity in our models is with market segmentation by for example income. However, individuals not only differ with respect to observed characteristics (e.g., income, sex, and race) but also with respect to the unobserved, but systematic, rules that they use for making judgments about choice alternatives. A common approach for accommodating both observed and unobserved heterogeneity is to use Mixed Logit (MXL) models, in which each respondent is assumed to follow his/her own choice rule. This is achieved by considering each weight as a random variable with an appropriate probability distribution. The model then estimates key parameters (e.g. mean and standard deviation) describing the distribution, allowing different individual weights to be attached to each attribute.

It can be shown (Hess et al, 2011) that allowing each taste parameter to take on different values across respondents induces some degree of correlation between the alternatives. However, it is still unclear if these flexibilities in the MXL model fully account for all correlation between the unobserved utilities of the alternatives. It is also possible that the model may actually be accounting for taste heterogeneity instead of inter-alternative correlations (Hess et al, 2005) leading to misleading conclusions. In Brownstone and Train (1999) they show how...
the mixed logit can be used to approximate any of the Gev models, however, such approximations are usually poor in practice due to simulation errors and are computational more expensive (Hensher and Green, 2001; Bhat, 2003).

In this paper we proposed the Mixed Nested Logit (MXNL) model which can be specified to simultaneously account for both taste-variations and inter-alternative correlations. This model was successfully implemented on actual data to measure drivers’ willingness to pay for using a toll road. The performance of this model was compared with the MNL, NL and MXL models in terms of model fits, substitution patterns and predictive power. The remainder of this article is organised as follows. In the following section, we give an overview of the theory, looking at MXNL models. Section 3 presents a summary of the empirical analysis conducted to explore the potential of MXNL models. Finally, we present the conclusions of the research in Section 4.

2. Methodology

2.1. The Utility Equation

The utility specification for an individual labeled $n$ ($n=1,2,...,N$), faced with a choice of choosing an alternative $j$ ($j = 1,2,...,J$) among $J$ number of alternatives is expressed as:

$$U_{nj} = V_{nj} + \varepsilon_{nj}$$

where:

- $U_{nj}$ is the overall utility for individual $n$ to alternative $j$
- $V_{nj}$ is the measured or observable utility for individual $n$ to alternative $j$
- $\varepsilon_{nj}$ is the unobservable utility or the error term for individual $n$ to alternative $j$

The measured part of the utility is the aspect of the utility that the analyst can observe or measure such as travel time or monetary cost. In this paper we assumed a linear specification of the measured utility as it simplifies the estimation process (Bhat and Koppelman, 2006) and it is also easier to interpret the estimated parameters.

$$V_{nj} = \sum_{k=1}^{K} \beta_{nk} X_{njk}$$

or could be expressed in a vector form as:

$$V_{nj} = \beta'_{n} X_{nj}$$

where:

- $X_{njk}$ = Represents individual $n$’s evaluation of alternative $j$ on attribute $k$ (e.g. travel time)
- $\beta_{nk}$ = "Shows the weight or importance that respondent $n$ attached to the corresponding attribute in the choice process and can be positive or negative depending on the attribute.

2.2. The Mixed Nested Logit (MXNL) Model

For a given vector of taste parameters $\beta$ the MXNL conditional probability of alternative $i$ for respondent $n$ is expressed as:

$$P_n(i/\beta) = \frac{\exp\left(\beta'_{n} X_{ni}/\mu_{m}\right) \prod_{j=1}^{J} \delta_{jm} \exp\left(\beta'_{n} X_{ni}/\mu_{m}\right)^{\mu_{s}}}{\sum_{j=1}^{J} \delta_{jm} \exp\left(\beta'_{n} X_{ni}/\mu_{m}\right) \prod_{m=1}^{M} \left(\sum_{j=1}^{J} \delta_{jm} \exp\left(\beta'_{n} X_{ni}/\mu_{s}\right)^{\mu_{s}}\right)^{\mu_{s}}}$$
Where $\delta_m$ is an indicator variable that equals 1 if alternative $i$ is assigned to nest $m$, and 0 otherwise. The parameter $\mu_m$ is called the structural or sensitivity parameter for nest $m$ and discrete choice theory suggests that this parameter should lie between 0 and 1 (Ortuzar, 1983; Daly and Zachary, 1978; Ben-Akiva and Lerman, 1985).

Since $\beta$ is unknown the unconditional probability is derived by integrating equation (1) over all possible values of $\beta$ weighted by its density function $f(\beta)$:

$$P_n(i) = \int \beta P_n(i/\beta) f(\beta/\Phi) d\beta$$

(5)

where $\Phi$ is a set of parameters describing the density function $f(\beta)$ and the corresponding log-likelihood function is given as:

$$LL(\Phi) = \sum_{n=1}^{N} \sum_{i=1}^{J} y_{n,i} \ln \left( \int \beta P_n(i/\beta) f(\beta/\Phi) d\beta \right)$$

(6)

where $y_{n,i}$ equals 1 if respondent $n$ chose alternative $i$ and zero otherwise.

The above integrals do not take a closed form, but they can be approximated through simulation. The resulting log-likelihood function becomes simulated log-likelihood:

$$SLL(\Phi) = \sum_{n=1}^{N} \sum_{i=1}^{J} y_{n,i} \ln \left( \frac{1}{R} \sum_{r=1}^{R} P_n(i/\beta_n^r) \right)$$

(7)

where the taste parameter $\beta$ can be expressed as:

$$\beta_n^r = b + \eta * N_n^r$$

(8)

being $b$ the mean and $\eta$ the standard deviation or the spread about the mean of the taste parameter of interest, and $N_n^r$ the $r$th ($r = 1, 2, \ldots R$) draw from the selected distribution for respondent $n$.

3. Empirical Analysis

3.1. Introduction

The data used for our empirical analysis comes from part of the survey data collected to measure the value of time during a toll road project in Nigeria in 2011. The overall approach involves a personal face-to-face interview of drivers at a roadside interview site who were using the main long distance spine route across Nigeria, some 1000 km long, which connected up major cities which were at least 1 hour apart. The roadside interview survey had questions about the trip they were currently making, their household and personal characteristics which was followed by the sp game. These respondents (drivers) were then asked to undertake a stated preference game, using their current journey as their reference trip. An orthogonal fractional factorial design was used with time and cost as the only attributes (apart from the existing road or toll road), each with two levels. The games were designed initially with prior knowledge about the values of time and were refined during the piloting stage. The time attribute levels ranged from 30 to 210 minutes. The cost attribute levels ranged from £3.20 to £8.0. All respondents were presented with various attribute levels within these ranges and asked to trade time, cost and whether to use the toll road or not. Respondents were asked to choose between their existing road and for the toll road alternative, they were asked to imagine their current trip being made via a toll road. The toll road was a new high quality dual carriageway highway while the existing road was generally a poor quality congested single carriageway road so we expected the Toll ASC to be positive (once the effect of the actual cost had been taken out).
The existing and toll alternatives each explored 4 levels of time and cost. These eight alternatives were presented to the respondents during the SP survey. The eight alternatives were grouped into two sets (existing road and toll road) and can therefore be considered as labelled. However, within the same set the alternatives are only differentiated by attributes and their levels describing them, so were unlabelled. This type of experiment can be considered as mixed labelled experiment.

The game was played in such a way that the respondent is first presented with two alternatives at a time (e.g., the alternative which was closest to their current trip which used the existing road versus the worse alternative using the existing road), and asked to state which alternative they would choose, and then, after they have made this choice, the winner (the chosen alternative) is placed in the top position and the loser (the non-chosen option) is placed below it. It is made to compete with any of the remaining alternatives for the second position. The winner for the second position then competes with the alternative at the first position else remains at the second position provided none of the remaining alternatives outperformed it. This is a process of successive pair wise comparisons. The process continues until the game is over, where the most preferred alternative is ranked first (ie at the top), followed by the second preferred alternative, and the least preferred ranked last (at the bottom). The resulting data constitute a ranking of the 8 alternatives that reflects the perceived utility that the respondent obtains from each alternative, with the alternative having the highest utility in the first position followed by the second best and so on. The ranking of the alternatives provided seven pseudo-observations for each respondent (see Train, 2009) where each alternative in turn is considered as a multinomial choice with all the alternatives below it as being the non-chosen alternatives.

A total of 281 drivers were intercepted and interviewed and 129 of them successfully completed the survey. As each respondents produced seven pseudo-observations, the total number of observations used in the estimation was 903 but 3 observations were rejected during the estimation process resulting in a total of 900 observations.

The explanatory variables used in the model fitting exercise included cost, and a toll constant. The toll constant was a constant for all the toll alternatives and was expected to reflect the perception of toll road by the respondents. All alternatives were variants of the current trip apart from the choice of existing versus toll road, so further ASC's were not considered. For the calibration of the various models discussed in this article, the estimation software Visual Choice was used. This estimation tool can be used for all types of closed-form as well as mixed GEV model structures and latent class models.

### 3.2. Model Fits Analysis

#### 3.2.1. The Multinomial Logit (MNL) Model

First, a simple MNL model was first fitted to the data; the estimation results for this model are reported in the first part of Table 3. As expected, the results showed negative marginal utilities for increases in travel time or travel cost. The estimated parameters are significant at 95% confidence interval and 54% goodness of fit. The toll constant is positive which is what we expect. It implies that all things being equal (e.g., equal time and cost) drivers will opt for the toll road as it is new and nicer than the untolled road. The MNL model produced a BIC (Bayesian Information Criterion, defined such that the smaller its value the better the model) of 1262, which is comparatively higher than the other models.

#### 3.2.2. The Nested Logit (NL) Model

To account for the potential existence of correlation between some of the alternatives, we grouped the toll routes into one nest called the toll nest and the rest into a non-toll nest. The resulting structural parameters were all significant at 95% level of confidence and within the expected range of 0 and 1. The results of this estimation are shown in the second part of Table 3. With this model structure, the structural parameters for non-toll and toll nests took the values 0.59 and 0.37 respectively, implying a relatively higher correlation between the unobserved utilities of the toll alternatives. Comparing the NL model with the MNL model, in terms of model fit, the results showed a very significant increase in Log-Likelihood (LL) by 51 units, with two additional parameters. This leads to a likelihood-ratio test value of 103, which has an associated chi-square p-value that is identical to zero (0.0) making the NL superior to the MNL.
3.2.3. The Mixed Logit (MXL) Model

The main issue of using MXL models is the choice of the number of random draws and the choice of distribution for the random parameters as they are not parameters to estimate in the model but user defined. For reasons of simplicity, a Normal distribution was used for the two coefficients. The use of bounded distributions may be more appropriate (Train and Sonnier, 2005, Hess, Bierlaire and Polak, 2004) however this sometimes leads to unacceptable large estimates (Hensher and Green, 2003). In the present application, the Normal distribution led to very good model performance. On the number of random draws we fitted the data on 0, 25, 100, 125, 150, and 200 Halton draws as shown in Table 1. From Table 1 the model appears to converge after 150 Halton draws. The 150 Halton draws was carried forward for further estimation and analysis.

The estimation results for the 150 draws are reported in the third part of Table 3. As expected, the results showed negative marginal utilities for increases in travel time or travel cost. The estimated parameters were 95% significant. In terms of model performance, with only two additional parameters, the MXL model led to significant improvements in the log-likelihood (LL) over MNL and the NL. The MXL model improved the log-likelihood (LL) over MNL and NL by 72 and 20 units respectively. The BIC statistic is also reduced by 130 and 41 for the MNL and NL models respectively.

Table 1: Mixed Logit (MXL) models with different number of Halton draws

<table>
<thead>
<tr>
<th>MXL</th>
<th>0</th>
<th>25</th>
<th>100</th>
<th>125</th>
<th>150</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Parameters</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>No of observations</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
</tr>
<tr>
<td>Null log likelihood</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
</tr>
<tr>
<td>Model log likelihood</td>
<td>-621</td>
<td>-569</td>
<td>-556</td>
<td>-556</td>
<td>-549</td>
<td>-553</td>
</tr>
<tr>
<td>Rho bar squared</td>
<td>54%</td>
<td>58%</td>
<td>59%</td>
<td>59%</td>
<td>60%</td>
<td>59%</td>
</tr>
<tr>
<td>BIC Statistic</td>
<td>1262</td>
<td>1171</td>
<td>1146</td>
<td>1145</td>
<td>1132</td>
<td>1139</td>
</tr>
</tbody>
</table>

3.2.4. The Mixed Nested Logit (MXNL) Model

Accounting for the existence of both inter-alternative correlation and taste heterogeneity, was done by grouping the toll routes into one nest called the toll nest and the rest into a non-toll nest and allowing the taste parameters to vary across respondents. Normal distributions were assumed for the two taste parameters (Time and cost) whilst the toll constant was non-random. Here again, we fitted the model using 0, 25, 100, 125, 150, and 200 Halton draws and based on the results in Table 2 we can comfortably conclude that 150 random draws seems to be satisfactory (see Table 2). The estimated parameters for the 150 Halton draws, are shown in column partition 4 of Table 3. The MXNL produced three set of parameters; means and the standard deviations of the taste parameters accounting for taste heterogeneity and the logsum parameters accounting for inter-alternative correlations. With this model structure, the structural parameters for non-toll and toll nests took the values 0.42 and 0.19 respectively, implying a significantly higher inter-alternative correlation within each nest than suggested by the Nested logit (NL) model, and support the NL claim that toll nest alternatives are more correlated. As expected, the results showed negative marginal utilities for increases in travel time or travel cost and implied value of time. The estimated coefficients are significant at the 95% confidence interval.

Comparing the MXNL model with the other models (MNL, NL, MXL), in terms of model fit, the results showed a very significant improvement in both model Log-Likelihood (LL) and the BIC values. It can be shown that all these three (MNL, NL and MXL) models are special cases of the MXNL model. Thus the use of the likelihood-ratio test may be appropriate to compare the MXNL and the other models as the authors are aware of the debate concerning the appropriateness of using this test statistic in comparing these models. The fitting of the MXNL model leads to a reduction in the model log-likelihood by 82, 31 and 10 over MNL, NL and MXL respectively. These results in likelihood-ratio tests values 165, 62 and 21 with associated chi-square p-values of approximately
0.0, 0.0 and 0.0 for the test between the MXNL and MNL, MXNL and NL, and MXNL and MXL respectively. Thus the MXNL is statistically superior to the MNL, NL and the MXNL models using the likelihood ratio test.

Table 2: Mixed Nested Logit (MXNL) models with different number of Halton draws

<table>
<thead>
<tr>
<th></th>
<th>MXNL</th>
<th>0</th>
<th>50</th>
<th>100</th>
<th>125</th>
<th>150</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Parameters</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>No of observations</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
</tr>
<tr>
<td>Null log likelihood</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
</tr>
<tr>
<td>Model log likelihood</td>
<td>-569</td>
<td>-542</td>
<td>-539</td>
<td>-539</td>
<td>-538</td>
<td>-541</td>
<td>-541</td>
</tr>
<tr>
<td>Rho bar squared</td>
<td>58%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>BIC Statistic</td>
<td>1173</td>
<td>1132</td>
<td>1126</td>
<td>1126</td>
<td>1124</td>
<td>1129</td>
<td>1129</td>
</tr>
</tbody>
</table>

Table 3: Estimated results for MNL, NL, MXL, MXNL models

<table>
<thead>
<tr>
<th>Variables</th>
<th>Multinomial (MNL)</th>
<th>Nested Logit (NL)</th>
<th>Mixed Logit (MXL)</th>
<th>Mixed Nested Logit (MXNL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVT (Min)</td>
<td>-0.1037</td>
<td>-0.0639</td>
<td>-0.5935</td>
<td>-0.1102</td>
</tr>
<tr>
<td>t-stats</td>
<td>30</td>
<td>18</td>
<td>7</td>
<td>24</td>
</tr>
<tr>
<td>Value of time (Pence/Min)</td>
<td>4.9</td>
<td>5.1</td>
<td>6.7</td>
<td>7.3</td>
</tr>
<tr>
<td>Cost (Pence)</td>
<td>-0.0211</td>
<td>-0.0124</td>
<td>-0.0889</td>
<td>-0.0151</td>
</tr>
<tr>
<td>t-stats</td>
<td>16</td>
<td>9</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Toll Constant</td>
<td>0.69</td>
<td>0.93</td>
<td>2.62</td>
<td>0.93</td>
</tr>
<tr>
<td>t-stats</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Relative to IVT</td>
<td>-6.7</td>
<td>-14.6</td>
<td>4.4</td>
<td>-8.5</td>
</tr>
<tr>
<td>Non Toll Logsum</td>
<td>0.59</td>
<td>0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stats</td>
<td>99</td>
<td>102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toll Nest Logsum</td>
<td>0.37</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stats</td>
<td>93</td>
<td>97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of Estimated Parameters</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>No of observations</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
</tr>
<tr>
<td>Null log likelihood</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
<td>-1365</td>
</tr>
<tr>
<td>Model log likelihood</td>
<td>-621</td>
<td>-569</td>
<td>-549</td>
<td>-538</td>
</tr>
<tr>
<td>Rho bar squared</td>
<td>54%</td>
<td>58%</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>BIC Statistic</td>
<td>1262</td>
<td>1173</td>
<td>1132</td>
<td>1124</td>
</tr>
</tbody>
</table>

3.3. Willingness To Pay Analysis

This section considers how the various models report the value of time (VOT). The MNL model produced an average VOT of about 4.90 pence per minute. The Nested Logit (NL) produced a slightly higher VOT of 5.10 pence per minute. Thus the VOT produced by the two homogeneous models are similar in magnitude. Also, both the MNL
and the NL reported only one value of time for all individuals, whilst the two mixed logit models reported both the means and standard deviations of values of time. The high and statistically significant standard deviations in both the MXL and the MXNL models support the need to account for taste heterogeneity among the individuals and that the assumption of homogeneous taste in the MNL and the NL appear to be invalid. The results therefore confirm the widely held view that different people place different values on travel time and/ or on travel cost. Also the mean VOT produced by the MXL and the MXNL models were higher than that of the MNL and the NL. For example, the mean VOT produced by the MXL model is higher by 1.5 pence per minute and 1.8 pence per minute than those of the NL and MNL respectively.

3.4. Substitution Patterns Analysis

Further analysis was carried out to ascertain the pattern of substitutions exhibited by these models. We investigated the shift in demand from two toll alternatives (alternatives 5 and 6) to the other alternatives by making them unavailable and re-running each model with the estimated parameters to forecast the new alternative shares. The alternative shares from the estimated models are presented under ‘Base’ in each model, whilst the new alternative shares resulting from making alternatives 5 and 6 not available are presented under ‘Forecast’ in table 4. Under the MNL model the highest alternative share from alternatives 5 and 6 not available is presented under ‘Forecast’ in table 4. Under the MNL model the highest alternative share from alternatives 5 and 6 has switched to alternative 7 with alternative 8 also receiving a lot. There was also a noticeable switch to alternative 2. Thus under the MNL model alternatives 7 and 8 accounted for only 58% of the switch from alternatives 5 and 6. This share has increased to 87% under the NL model, with alternative 2 (in the non-toll nest) now having a less significant share, indicating that the existence of inter-alternative correlation has not been properly accounted for by the MNL model. Under the MXL model the share for the remaining two toll-alternatives has reduced to 74%, an indication that a ‘chunk’ of the existence of inter-alternative correlation has been accounted for by this model (but at all). Therefore the MXL model seems to have accounted for some of the inter-alternative correlations that the NL model was accounting for.

Comparing the MXNL with NL shows that the remaining two alternatives in the toll nest accounts for 88% of the switch under the MXNL model which is similar to the NL’s 87%. This indicates that the two models may be exhibiting similar substitution patterns which is expected because the strength of the NL is in its ability to effectively explain substitution patterns. In contrast the mixed logit is poor at explaining substitution patterns (but good at explaining taste heterogeneity). Clearly, the MXL accounted for some of the inter-alternative correlations but may have masked the rest. This analysis shows the important gains in model fit and predictive power in using the more complex model structures considered here.

Table 4: Investigating substitution patterns

<table>
<thead>
<tr>
<th>MNL Base</th>
<th>Forecast</th>
<th>Diff</th>
<th>Share</th>
<th>NL Base</th>
<th>Forecast</th>
<th>Diff</th>
<th>Share</th>
<th>MXL Base</th>
<th>Forecast</th>
<th>Diff</th>
<th>Share</th>
<th>MXNL Base</th>
<th>Forecast</th>
<th>Diff</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 109</td>
<td>108</td>
<td>-1</td>
<td>0%</td>
<td>113</td>
<td>109</td>
<td>-4</td>
<td>-2%</td>
<td>120</td>
<td>123</td>
<td>3</td>
<td>1%</td>
<td>118</td>
<td>112</td>
<td>-5</td>
<td>-2%</td>
</tr>
<tr>
<td>2 125</td>
<td>201</td>
<td>76</td>
<td>33%</td>
<td>125</td>
<td>147</td>
<td>23</td>
<td>10%</td>
<td>140</td>
<td>183</td>
<td>43</td>
<td>20%</td>
<td>124</td>
<td>150</td>
<td>26</td>
<td>11%</td>
</tr>
<tr>
<td>3 133</td>
<td>137</td>
<td>4</td>
<td>2%</td>
<td>127</td>
<td>133</td>
<td>7</td>
<td>3%</td>
<td>140</td>
<td>147</td>
<td>8</td>
<td>3%</td>
<td>132</td>
<td>136</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>4 33</td>
<td>50</td>
<td>18</td>
<td>8%</td>
<td>35</td>
<td>41</td>
<td>5</td>
<td>2%</td>
<td>17</td>
<td>20</td>
<td>3</td>
<td>1%</td>
<td>26</td>
<td>29</td>
<td>3</td>
<td>1%</td>
</tr>
</tbody>
</table>

N toll Nest Share | 42% | 13% | 26% | 12%

Toll Nest Share | 58% | 87% | 74% | 88%

Total | 900 | 900 | 234 | 900 | 900 | 235 | 900 | 900 | 217 | 900 | 900 | 233 | 968 |
3.5. Predictive Power Analysis

We also investigated the similarity of the models in predicting the observed choices by regressing the probabilities of the chosen alternatives of the MXNL against those of MNL, NL and MXL. In general all the models appear to be predicting the observed alternative choices closely with some slight differences for some individuals. The regressions in Fig. 1, 2 and 3 suggest that the choice probabilities of the NL accounted for most of the variations in the choice probabilities of the MXNL with R-squared of 92.7% (Fig. 1). This was followed by the MXL with R-squared of 91.5% (Fig. 2) and then the MNL with R-squared of 83.8% (Fig 3). These differences in variability may suggest that the effects of inter-alternative correlation could be stronger than that of the taste heterogeneity in the choice process.

Fig. 1: Predictive power of the models: NL vs. MXNL

Fig. 2: Predictive power of the models: MXL vs. MXNL
4. Conclusion

This paper investigated the model formed when mixed logit is combined with nested logit (MXNL) to simultaneously account for taste heterogeneity and inter-alternative correlations using a real Stated Preference (SP) dataset from toll route choice experiments conducted during a recent toll route study. The performance of the MXNL model was compared with the Multinomial logit (MNL), Nested logit (NL) and Mixed logit (MXL) in terms of model fit, substitution patterns, predictive power and evaluation of implied willingness to pay. Among the GEV models, the NL model with 2 nests (toll nest consisting of toll route alternatives and non-toll nest consisting of non-toll route alternatives) was shown to be statistically better than the MNL. However, both the NL and MNL models were shown to be unsuitable for properly accounting for the variations in taste across respondents. The existence in taste variations across the respondents were shown by the MXL and the MXNL models to be very significant. This was demonstrated by the highly significant standard deviations of the values of time (VOT) reported by the MXL and the MXNL models. The homogenous assumption of the MNL and the NL models resulted in lower than expected values of time of these models.

The paper also demonstrated how the MNL model failed to exhibit the correct substitution pattern when one or more alternatives become unavailable. It failed to properly account for the fact that the toll road alternatives have similar unobserved factors and hence better substitute for each other than the non-toll alternatives. This phenomenon was revealed by the NL and the MXNL models. The MXL model was found to have accounted for some of these inter-alternative correlations and may have masked the rest.

We also investigated the similarity of the models in predicting the observed choices by regressing the probabilities of the chosen alternatives of the MXNL against those of MNL, NL and MXL and found that all the models were predicting the observed alternative choices closely. The regressing analysis suggests that the effects of inter-alternative correlation may be stronger than that of the taste heterogeneity in the choice process.

The paper also shows the risk of accounting for only one of these phenomena as one could mask the presence of the other. This study supports the work by Hess et al. (2005) and Teye et al. (2013) advocating the use of model structures which simultaneously account for both taste heterogeneity and inter-alternative correlation for situations where the nature of the error-structure is not clear a priori.

This paper illustrates how these additional tools can better explain the value of time and substitution patterns thereby helping analysts improve their toll road forecasts.

References


Davidson, P. W and Teye, Collins, Cully, R. Understanding the potential role of the latent mix model comprising a combination of latent class and mixed logit European Transport Conference, Glasgow, 2012.
Daniel McFadden (2000), Disaggregate Behavioral Travel Demand’s RUM Side, A 30-Year Retrospective Department of Economics, University of California, Berkeley, March 2000


