Transportmetrica, Vol. 4, No. 2 (2008), 155-177

DISCRETE CHOICE MODELING OF COMBINED MODE AND DEPARTURE TIME

SHAMAS BAJWA[1], SHLOMO BEKHOR[2], MASAO KUWAHARA[3] AND EDWARD CHUNG[4]

Received 15 January 2007; received in revised form 27 September 2007; accepted 8 January 2008

Typical daily decision-making process of individuals regarding use of transport system involves mainly three types of decisions: mode choice, departure time choice and route choice. This paper focuses on the mode and departure time choice processes and studies different model specifications for a combined mode and departure time choice model. The paper compares different sets of explanatory variables as well as different model structures to capture the correlation among alternatives and taste variations among the commuters. The main hypothesis tested in this paper is that departure time alternatives are also correlated by the amount of delay. Correlation among different alternatives is confirmed by analyzing different nesting structures as well as error component formulations. Random coefficient logit models confirm the presence of the random taste heterogeneity across commuters. Mixed nested logit models are estimated to jointly account for the random taste heterogeneity and the correlation among different alternatives. Results indicate that accounting for the random taste heterogeneity as well as inter-alternative correlation improves the model performance.

KEYWORDS: Departure time choice, mode choice, nested logit, cross-nested logit, error component logit, mixed logit, value of time, value of schedule delay

1. INTRODUCTION

Advent of the advanced transport control and communication technologies has made it possible to implement time-varying demand management policies for example, time-varying road pricing. It is necessary to develop the departure time and mode choice behavioural models to assess the impact of these policies.

This paper discusses the development of a combined mode and departure time choice model for morning commuters. Typical daily decision-making process of individuals regarding use of transport system involves mainly three types of decisions: departure time choice, mode choice and route choice. In this paper, we focus on the first two components of the choice behavior.

Most commuters have preferred arrival times at their destinations due to constraints of work (school) starting times. These arrival times are generally concentrated in a short time period indicating peak demand during rush hours. Time-dependent demand management policies attempt to spread the peak demand on a longer time period, by providing commuters a trade-off between arriving early/late at destination than the preferred arrival time and thus spending less time in congestion (cost) or arriving on-time but spending more time in congestion (cost). Hence, the parameter of interest is not only the value of travel time savings which is usually estimated in traditional mode choice models but also the value of early/late schedule delay to establish the trade-off among different alternative departure/arrival times available to commuters.

In this paper, our main focus is the departure time choice behavior of road traffic network users. However, it is also known that mode choice always stands as a viable option if alternative public transport modes provide extensive spatial coverage with high

[1] School of Civil, Environmental and Chemical Engineering, RMIT University, Melbourne, Australia. Corresponding author (E-mail: shamas.bajwa@rmit.edu.au).

[2] Transportation Research Institute, Technion – Israel Institute of Technology, Haifa, Israel.

[3] Institute of Industrial Science, The University of Tokyo, Tokyo, Japan.

[4] Institute of Industrial Science, The University of Tokyo, Tokyo, Japan.

frequency which is usual for metropolitan areas. Need to include mode choice in commuters' choice behavior models is further highlighted if we want to test some travel demand management strategies e.g. road pricing which may be prohibitively expensive for some section of the commuters. Hence, mode choice can be considered to provide a viable alternative to travel. Mode choice can also be considered as a stand-in for elastic demand in the traffic networks by assuming a constant travel time between origin and destination as well as a fixed cost and schedules will keep its utility constant, giving it the role of a null option in the elastic demand models.

Important considerations for behaviour modelling include appropriate specification of the utility functions associated with all alternatives. Users' choices depend not only on their socio-economic characteristics but also on the level of service attributes of the network used for commuting. An appropriate utility function specification should include a mix of socioeconomic characteristics and level of service attributes explaining the maximum variance among the user's behaviour. However, it is still possible that the taste variations among the commuters are not captured due to some unknown characteristics or measurement errors. These random (unmeasured and/or non-quantifiable) taste variations among commuters can significantly affect the performance of the model, and should be taken into account. The progress in the field of the mixed logit models allows the estimation of such taste variations by assuming a distribution of parameters over the commuters' population instead of a single value. In addition, flexibility of the model structure allows capturing correlation among different alternatives. In the case of discrete choice models of departure time, the consecutive departure time intervals can be highly correlated due to the continuity of the underlying variable (i.e. time) and inability of the commuters to distinguish between the close by alternative departure times.

In this paper, our aim is to identify appropriate utility specifications and capture the random taste heterogeneity across the commuters by using the mixed logit class of models. In addition, correlation among alternatives is investigated by specifying different nesting structures as well as error components. Different models are compared and discussed to highlight their relative strengths and weaknesses.

Two main contributions can be outlined in this paper as follows. First, it provides a comprehensive comparative study of different model structures based on the same dataset, by showing the relative improvement over successive flexible models. Second, the paper shows that departure time alternatives are also correlated by the amount of delay, by examining different correlation structures among alternatives.

The remaining of this paper is organised as follows: Next section describes in brief previous attempts to model departure time choice behaviour as well as a brief introduction to the mixed logit and nested logit models. Section 3 details the methodology used in this study, including the survey design, data collection and model specifications. Section 4 shows the results of the estimated models and discusses the improvements, relative merits and demerits as well as problems encountered in different specifications. Finally, section 5 concludes the paper with a summary of the findings and outlines the future research directions.

2. LITERATURE SURVEY

Nearly all the existing models about departure time choice are based on the trade-off between early, late or on-time arrival first proposed by Vickrey (1969). Assuming a single bottleneck on the route, the departure time choice is analyzed by evaluating the trade-off between waiting in queue and arriving on-time or starting early/late and arriving earlier/later than a preferred arrival time. Hendrickson and Kocur (1981) reformulated the same problem and elaborated the analysis using the queuing theory notation. Henderson (1974), Hendrickson et al. (1981), Hurdle (1981), Fargier (1981) also independently solved the departure time choice problem for a single bottleneck case (i.e. route choice is not considered). Kuwahara (1985) and Kuwahara and Newell (1987) extended the analysis of departure time choice in a network to a many-to-one origin destination network, where each commuter passes only one bottleneck. de Palma et al. (1983) used a similar construction of one origin and destination with a single bottleneck as used by Vickrey (1969) and others for a stochastic equilibrium model of departure time choice. All models above consider the interaction among supply and demand and provide equilibrium solution. Same trade-off principle is used in all these models and is also employed for discrete choice model estimation.

Most discrete departure or arrival time models developed are based on the multinomial logit (MNL) model: for example, Small (1982) used the data collected from the car commuters in the San Francisco bay area to model the arrival time choice, Hendrickson and Plank (1984) estimated a combined mode and departure time choice model. Abkowitz (1981) also used the same data set as used by Small (1982) including additional socio-demographic variables as determinants of commuter's departure time choice behavior. Chin (1990) modeled the departure time choice of morning commuters using the data collected in Singapore. Shimizu and Yai (1999) carried out a survey to gauge the reaction of commuters to a variable peak period toll in Tokyo Metropolitan area. The departure time choice of the users was modeled as a discrete logit choice in half an hour intervals including the shift to public transport or alternative un-tolled route as a choice at the same level as departure time choice.

Use of MNL models ignores any correlation among the consecutive discrete departure time intervals. If the departure time interval becomes small, it is for the decision-makers to distinguish between the adjacent time intervals resulting in a higher correlation between alternatives. Small (1987) proposed an OGEV model for the departure time choice which has a more flexible correlation structure than MNL model by allowing for the correlation parameter to exist for pairs of alternatives which depends on the distance among the alternatives based on time-of-day natural ordering. The number of correlated alternatives needs to be specified before-hand. Bhat (1998a) used MNL for modeling mode choice and an ordered generalized extreme value (OGEV) for departure time choice. The proposed MNL-OGEV model was applied to data obtained from the 1990 San Francisco Bay area travel survey data and was found to perform better than the MNL and nested logit models. Results indicate that the MNL and nested logit models lead to biased level-of-service estimates and to inappropriate policy evaluations of transportation control measures. Polak and Jones (1994) used a nested logit structure to model the departure time choice in a tour based context.

Cross-nested logit models extend and generalize the correlation structure among the alternatives. Instead of each alternative belonging to a single nest in nested logit models, cross-nested models allow alternatives to belong to more than one nest thus resulting in a flexible correlation structure. These models define the share of each alternative belonging to different nests. Recently similar flexible correlation structures have been developed and used as cross-nested, generalized nested and paired combinatorial logit by many researchers (Vovsha, 1997; Vovsha and Bekhor, 1998; Koppelman and Wen, 2000; Wen and Koppelman, 2001; Papola, 2004).

Mixed logit is a highly flexible model that can approximate any random utility model (McFadden and Train, 2000). It obviates the three limitations of standard logit model by allowing for random taste variations, unrestricted substitution patterns, and correlation among unobserved factors over time. Unlike probit, mixed logit is not restricted to normal distributions, have straightforward derivation and choice probabilities can be simulated easily (Train, 2003). The first application of mixed logit was apparently the automobile demand models created by Boyd and Mellman (1980) and Cardell and Dunbar (1980).

Bhat (1998b) used mixed multinomial logit model for analysis of travel mode and departure time choice for home-based social–recreational trips using data drawn from the 1990 San Francisco Bay Area household survey. The empirical results highlighted the need to capture unobserved attributes for mode as well as departure time which not only improved model fitness but also resulted in realistic evaluations of transportation control measures. de Jong et al. (2003) also developed an error component logit model for the joint choice of time-of-day and mode using stated preference data for car and train travelers in The Netherlands. The results indicate the time-of-day choice is sensitive to the peak travel time and cost. A different approach to model departure time has been to use continuous time models instead of discrete time intervals (Wang, 1996; Bhat and Steed, 2002).

In most researches, the multinomial logit is used as the kernel for the mixed logit model. Recently, Hess et al. (2004) have applied different model structures such as mixed nested logit model and mixed cross-nested logit model for the mode choice. They proposed these modeling structures to capture the effect of the random taste heterogeneity as well as the inter-alternative correlation. This study showed that use of mixed GEV models improves the performance compared to basic models.

Review of the past literature reveals that different model structures have been applied for disaggregate departure time choice models but a comprehensive study comparing different model structures estimated for the same data set has not been carried out. Furthermore, most of the existing departure time choice models are developed solely for departure time choice of the commuters while it may be reasonable to assume that commuters decide their mode and departure time concurrently. In this research, a comprehensive combined departure time and mode choice modeling study is carried out in order to compare the relative pros and cons of different modeling structures.

3. METHODOLOGY

Discrete choice models are used to replicate the choices made by decision-makers (i.e. commuters) from a discrete number of alternatives which constitute the choice set depending on the availability. The methodology used to specify and estimate combined departure time and mode choice models are described in the subsequent sections, organised as follows. The first section briefly presents the model structures used in this paper. Data describing users' behaviour is presented next. After that choice set definition is described. The last subsection presents the specification of the utility functions and correlation structures.

3.1 General model structure

The usual form of the utility function is

$$U_{ij} = V_{ij} + \varepsilon_{ij}\,, \tag{1}$$

where U_{ij} is the utility of the individual *j* for alternative $i \in C_j$, V_{ij} is the deterministic part of the utility of the alternative *i* for individual *j*, and ε_{ij} is the random component of the utility of the alternative *i* for individual *j*.

$$V_{ij} = f\left(\beta, x_{ij}\right), \tag{2}$$

where x_{ij} is a vector representing the attributes of an alternative *i* as well as the socio-economic characteristics of the decision-maker *j*, and β is vector of coefficients which needs to be estimated from the data. Depending on the assumed form of the random component of the utility, different models can be developed.

In the random coefficient logit model, some of the elements of the vector β described in equation (2) are declared as random variables to capture the taste heterogeneity in the population. In the error component logit model, instead of some elements of vector β corresponding to the vector x_{ij} being random, separate random term is introduced in the utility function,

$$V_{ij} = f\left(\beta, x_{ij}\right) + \xi\,, \tag{3}$$

where, ξ is a random disturbance, generally assumed to follow a multivariate normal distribution with a mean zero and covariance matrix Ω, where Ω is generally assumed to be diagonal (Walker, 2001).

3.2 Survey design and data collection

A stated preference survey was conducted in Tokyo Metropolitan Area to elicit the responses of the users corresponding to different hypothetical scenarios specifying different departure times as well as travel times and costs to reach destination at a preferred arrival time. Households were randomly selected and data about the primary morning commuter was collected using a mail-back survey. A total of 1324 valid responses were used for model estimation.

The problem posed to the users was as follows: given a specific arrival time, commuters have different departure options from home; they can select the mode as well as the departure time. Commuters are assumed to choose their mode between car and rail (the only public transport mode presented in the survey). Departure time was modeled in 15 minute intervals and this option was only available if commuters chose car as mode. Car commuters can trade-off between arriving early or late with less congestion (i.e. shorter travel time) or arriving on-time with higher congestion (i.e. longer travel time) on the road. Different levels of toll were also introduced. This allowed us to measure the trade-off among the monetary costs, travel times and schedule delay penalties. It was assumed that rail users can reach their destination without any schedule delay and with a fixed cost. This assumption is quite reasonable due to the high frequency of trains in the region. The only aspect of the public transport not accounted for in this study is the congestion levels inside the train as the commuters can choose a different departure time from their home while using public transport to commute in order to avoid severe congestion inside the trains but anecdotal evidence suggests to the contrary.

The questionnaire presented to the users consisted of two sections: in first section, socio-economic characteristics of the household as well as the commuter were collected.

The collected information consists of the household type (single, couple, couple with children etc.), dwelling type and number of cars in household. Personal information collected includes the personal characteristics such as age, gender, income, work location (post-code) and information about morning commute on a typical day. In second part, stated preference scenarios were presented and users were asked to choose departure time as well as mode. Cost and travel time of the rail were fixed at 500 yen (~$4) and 60 minutes respectively while for the car, cost was varied between three levels of 500, 700 and 1000 yen and the travel time was varied at five levels from 40 to 60 minutes at 5 minute intervals. The early and late arrival delay was automatically deducted from the interaction of the departure time, travel time and preferred arrival time at the destination.

To ensure the statistically efficient information retrieval from the collected data while not cognitively burdening the users excessively, a standard statistical experiment design procedure named factorial design was used in this study. A fractional factorial design, which can cater for the main effects as well as some first-order interactions among the attribute levels of different alternatives, was used. Fractional factorial design means loss of some statistical efficiency by ignoring second and higher order effects i.e. interactions among two or more than two attributes. But the loss of information is not critical as it has been shown that more than 80% of the information is explained by main effects, 15% by the first order effects and remaining 5% by the second and higher order effects (Louviere et al., 2000). Hence, proper fractional factorial design can be used to design experiments with over 95% statistical efficiency.

3.3 Choice set definition

Based on the information collected from the survey respondents in the stated preference survey, the selected alternatives as well as the alternative set presented to the subjects in each scenario can be aggregated as proposed by Cascetta and Papola (2003). The number of choices presented in each scenario were limited to three to avoid the cognitive load to the survey respondents where one alternative was always rail indicating its availability to all the users independent of their location. Remaining two alternatives were different departure time options using car. As each respondent was presented with a maximum of either two early arrivals or two late arrivals due to rail being constrained to on-time arrival, hence the alternatives can be aggregated into following six alternatives available to each user without any loss of information in data:

- Earliest Early Arrival Car (EEA)
- Latest Early Arrival Car (LEA)
- On-time Arrival Car (OT)
- Earliest Late Arrival Car (ELA)
- Latest Late Arrival Car (LLA)
- Rail (RL)

This aggregation can be justified because many of the alternatives may never be chosen in the sample because of its size and consequently not included in the final choice set. No departure time option is available for the rail because of its frequency. The frequency of rail in Tokyo area in the morning is high and all the rail users can choose rail which allows them to reach their destination without any schedule delay. Table 1 shows a summary of the choices and availabilities of each alternative in the sample.

TABLE 1: Choices and availabilities of alternatives in the sample

Alternatives	Choices	Availabilities
Earliest Early Arrival Car (EEA)	28	345
Latest Early Arrival Car (LEA)	332	1241
On-time Arrival Car (OT)	24	130
Earliest Late Arrival Car (ELA)	50	880
Latest Late Arrival Car (LLA)	4	52
Rail (RL)	886	1324
Total	1324	

3.4 Model structure specifications and selected attributes

As the correlation among the alternatives is not known in advance, we need to hypothesize and test different correlation structures to identify the best fitting and explanatory model. The basic structure tested is an MNL model assuming that no correlation exists between any of the alternatives. The nesting structure as well as covariance matrix form for this formulation is as shown in Figure 1.

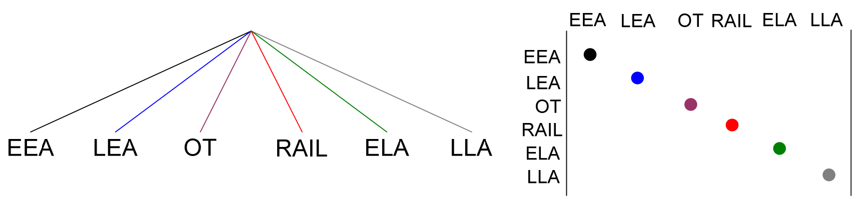


FIGURE 1: Multinomial logit correlation structure and relevant covariance matrix

Figure 2 shows a nested logit model structure in which the alternatives are grouped together by mode i.e. rail is a separate nest while the departure time options corresponding to car are grouped together in a single nest. The covariance matrix corresponding to this structure is also shown in Figure 2 indicating that alternatives EEA, LEA, OT, ELA and LLA are correlated with each other while the rail is not.

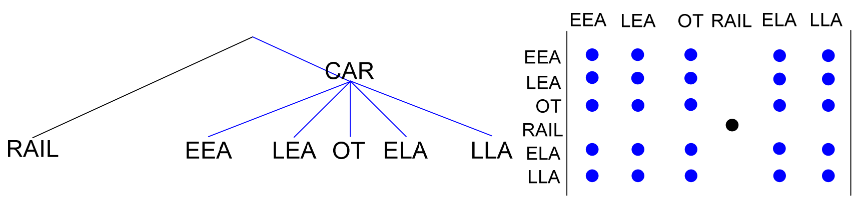


FIGURE 2: Nested logit (mode based) correlation structure and relevant covariance matrix

Figure 3 shows another nesting structure in which the alternatives are assigned to three nests. Rail is in a separate nest i.e. is not correlated with any other alternative while the departure times using car resulting in early or on-time arrival at the destination are grouped in one nest while the alternatives depicting the departure by car for late arrivals are grouped together in one nest. The corresponding covariance matrix structure is as shown in the Figure 3.

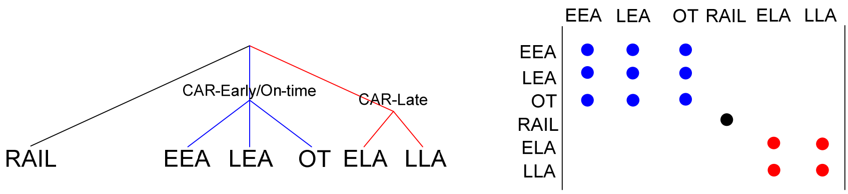


FIGURE 3: Nested logit (3nests-a) correlation structure and relevant covariance matrix

Figure 4 shows another nesting structure indicating that alternatives are grouped together based on the arrival time at the destination irrespective of the mode. Three nests are formed each corresponding to early, on-time and late arrivals. Each nest has two alternatives. The corresponding covariance matrix is also shown in the Figure 4.

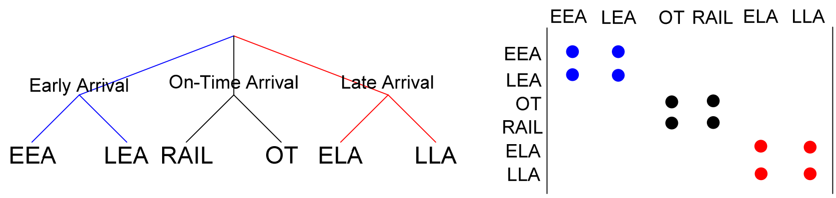


FIGURE 4: Nested logit (3nests-b) correlation structure and relevant covariance matrix

Figure 5 shows the nesting structure with four nests, one corresponding to the early arrivals using cars, another corresponding to late arrivals using car while the remaining two indicate the on-time arrivals at the destination but using different modes. The corresponding covariance matrix is also shown in the Figure 5.

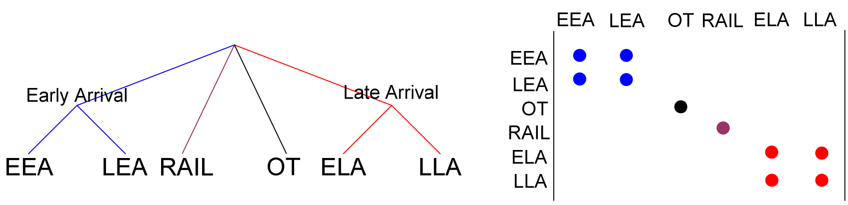


FIGURE 5: Nested logit (4nests) correlation structure and relevant covariance matrix

The attributes used in the modeling stage can be divided into two distinct groups: level of service attributes and personal characteristics of each individual user. Different attributes tried for the model estimations are defined as follows:

$\beta_{\text{travel time}}$: Coefficient for travel time, where travel time is given in minutes

β_{cost} : Coefficient for cost of travel, where cost is in yen

$\beta_{\text{early arrival}}$: Coefficient of early arrival penalty, where early arrival penalty is calculated based on the departure time as well as preferred arrival time at the destination, in case of random coefficient model, this represents the mean of the coefficient distribution

$\sigma_{\text{early arrival}}$: Variance of the early arrival penalty in random coefficient model

$\beta_{\text{late arrival}}$: Coefficient of late arrival penalty, where late arrival penalty is calculated based on the departure time as well as preferred arrival time at the destination, in case of random coefficient model, this represents the mean of the coefficient distribution

$\sigma_{late\ arrival}$: Variance of the late arrival penalty in random coefficient model

$\beta_{car\ availability}$: Coefficient of car availability, where Car Availability is a dummy variable equal to 1 if commuter owns a car; 0 otherwise

$\beta_{old\ age}$: Coefficient representing the effect of old age, where old age is a dummy variable equal to 1 if commuter is more than 70 years old and 0 otherwise

$\beta_{high\ income}$: Coefficient representing the effect of high income, where high income is a dummy variable equal to 1 if commuter's annual income is more than 15 million yen and 0 otherwise

β_{young} : Coefficient representing the behaviour of young people, where young is a dummy variable if commuter is younger than 30 years of age or is a student and 0 otherwise

$\beta_{work\ in\ suburbs}$: Coefficient representing the effect if the commuter's work place is not in central Tokyo, where work in suburbs is a dummy variable equal to 1 if the commuter's work place is in suburbs and 0 otherwise; in case of random coefficient model, this represents the mean of the coefficient distribution

$\sigma_{work\ in\ suburbs}$: Variance of the work in suburbs coefficient in random coefficient model

ξ_{car} : Random error component constrained to be same for all the alternatives using car as a mode

ξ_{rail} : Random error component constrained to be same for the alternative using rail as a mode

$\xi_{on\text{-}time}$: Random error component constrained to be same for the alternative resulting in on-time arrival

$\xi_{early\ arrival}$: Random error component constrained to be same for the alternative resulting in early arrival

$\xi_{late\ arrival}$: Random error component constrained to be same for the alternative resulting in late arrival

$\xi_{early\ arrival/on\text{-}time}$: Random error component constrained to be same for the alternative resulting in early or on-time arrival

4. MODEL ESTIMATION RESULTS

Maximum likelihood and simulated maximum likelihood methods were used for estimating the parameters of the closed-form GEV and mixed GEV models respectively. These methods try to maximize the log-likelihood function. As stated earlier, the data collected in a survey of the morning commuters in the Tokyo Metropolitan area is used in this study. Estimation software BIOGEME is used for model estimations (Bierlaire, 2005)

4.1 MNL models

MNL model as depicted in Figure 1 is the first estimated model structure for combined mode and departure time. Different utility specifications were tested to find the best possible utility function explaining the maximum variance in the data. Two alternative utility specifications are mentioned below. One only includes the level of service (LOS) attributes while other also includes the personal characteristics of the commuters. The model using the personal characteristics in addition to the LOS attributes shows a marked improvement over the LOS only model. The log-likelihood value increases by about 70 points by addition of the 5 personal characteristics. The personal characteristics chosen for the inclusion in the model are those which have been found to be significant during different trial estimations. The LOS only model is good for the network-wide applications where the detailed data about the personal characteristics of the commuters is not available. All the utility functions are linear in attributes as well as parameters. The estimation results of the MNL models are reported in the Table 2 with t-statistics shown in brackets. All the parameters are found to be significant at a level greater than 95%. The value of travel time savings is about 38 yen/min or about 2300 yen/hour (about US$20). The value of early and late arrival penalty is about 22 yen/min and 110 yen/min respectively which is about 60% and 300% of the value of travel time savings. These values are quite similar to what have been reported elsewhere in literature for different geographic locations.

TABLE 2: Estimation results of MNL model of Figure 1

Coefficients	Level of service attributes		Level of service + socio-demographic attributes	
ASC_{Rail}	0.7040	(5.13)	1.8450	(9.38)
$\beta_{\text{travel time}}$	-0.0300	(-3.84)	-0.0314	(-3.91)
β_{cost}	-0.0008	(-3.15)	-0.0008	(-3.21)
$\beta_{\text{early arrival}}$	-0.0167	(-7.03)	-0.0180	(-7.28)
$\beta_{\text{late arrival}}$	-0.0844	(-7.78)	-0.0883	(-7.92)
$\beta_{\text{car availability}}$			0.9535	(6.64)
$\beta_{\text{old age}}$			1.3211	(2.48)
$\beta_{\text{high income}}$			1.1410	(4.20)
β_{young}			-0.5656	(-2.91)
$\beta_{\text{work in suburbs}}$			-1.0320	(-7.90)
No. of observations	1324		1324	
No. of parameters	5		10	
Null-log likelihood	-1454.56		-1454.56	
Final-log likelihood	-1052.76		-982.91	
Rho-Squared	0.276		0.324	
Rho-Squared bar	0.273		0.317	
VTTS(yen/min)	38.0		38.5	
VEAP(yen/min)	21.2		22.0	
VLAP(yen/min)	107.1		108.2	

VTTS = Value of Travel Time Savings
VEAP = Value of Early Arrival Penalty
VLAP = Value of Late Arrival Penalty

The utility functions for the LOS only model is

$$V_{\text{car,EEA}} = \beta_{\text{travel time}} TT_{\text{EEA}} + \beta_{\text{cost}} COST_{\text{EEA}} + \beta_{\text{early arrival}} EA_{\text{EEA}} + \beta_{\text{late arrival}} LA_{\text{EEA}}$$

$$V_{\text{car,LEA}} = \beta_{\text{travel time}} TT_{\text{LEA}} + \beta_{\text{cost}} COST_{\text{LEA}} + \beta_{\text{early arrival}} EA_{\text{LEA}} + \beta_{\text{late arrival}} LA_{\text{LEA}}$$

$$V_{\text{car,OT}} = \beta_{\text{travel time}} TT_{\text{OT}} + \beta_{\text{cost}} COST_{\text{OT}} + \beta_{\text{early arrival}} EA_{\text{OT}} + \beta_{\text{late arrival}} LA_{\text{OT}}$$

$$V_{\text{car,ELA}} = \beta_{\text{travel time}} TT_{\text{ELA}} + \beta_{\text{cost}} COST_{\text{ELA}} + \beta_{\text{early arrival}} EA_{\text{ELA}} + \beta_{\text{late arrival}} LA_{\text{ELA}}$$

$$V_{\text{car,LLA}} = \beta_{\text{travel time}} TT_{\text{LLA}} + \beta_{\text{cost}} COST_{\text{LLA}} + \beta_{\text{early arrival}} EA_{\text{LLA}} + \beta_{\text{late arrival}} LA_{\text{LLA}}$$

$$V_{\text{Rail}} = ASC_{\text{Rail}} + \beta_{\text{travel time}} TT_{\text{Rail}} + \beta_{\text{cost}} COST_{\text{Rail}}$$

While for the model including level of service as well as personal attributes, the best possible utility function specification is

$$
\begin{aligned}
V_{\text{car,EEA}} &= \beta_{\text{travel time}} TT_{\text{EEA}} + \beta_{\text{cost}} COST_{\text{EEA}} + \beta_{\text{early arrival}} EA_{\text{EEA}} + \beta_{\text{late arrival}} LA_{\text{EEA}} \\
&\quad + \beta_{\text{car availability}} CarAvailability + \beta_{\text{high income}} HighIncome \\
V_{\text{car,LEA}} &= \beta_{\text{travel time}} TT_{\text{LEA}} + \beta_{\text{cost}} COST_{\text{LEA}} + \beta_{\text{early arrival}} EA_{\text{LEA}} + \beta_{\text{late arrival}} LA_{\text{LEA}} \\
&\quad + \beta_{\text{car availability}} CarAvailability + \beta_{\text{high income}} HighIncome \\
V_{\text{car,OT}} &= \beta_{\text{travel time}} TT_{\text{OT}} + \beta_{\text{cost}} COST_{\text{OT}} + \beta_{\text{early arrival}} EA_{\text{OT}} + \beta_{\text{late arrival}} LA_{\text{OT}} \\
&\quad + \beta_{\text{car availability}} CarAvailability + \beta_{\text{high income}} HighIncome \\
V_{\text{car,ELA}} &= \beta_{\text{travel time}} TT_{\text{ELA}} + \beta_{\text{cost}} COST_{\text{ELA}} + \beta_{\text{early arrival}} EA_{\text{ELA}} + \beta_{\text{late arrival}} LA_{\text{ELA}} \\
&\quad + \beta_{\text{car availability}} CarAvailability + \beta_{\text{high income}} HighIncome \\
V_{\text{car,LLA}} &= \beta_{\text{travel time}} TT_{\text{LLA}} + \beta_{\text{cost}} COST_{\text{LLA}} + \beta_{\text{early arrival}} EA_{\text{LLA}} + \beta_{\text{late arrival}} LA_{\text{LLA}} \\
&\quad + \beta_{\text{car availability}} CarAvailability + \beta_{\text{high income}} HighIncome \\
V_{\text{Rail}} &= ASC_{\text{Rail}} + \beta_{\text{travel time}} TT_{\text{Rail}} + \beta_{\text{cost}} COST_{\text{Rail}} + \beta_{\text{old age}} OldAge \\
&\quad + \beta_{\text{young}} Young + \beta_{\text{work in suburbs}} WorkInSuburbs
\end{aligned}
$$

A positive alternative specific constant for rail indicates an inherent preference to choose rail over other mode which is quite understandable owing to the chronic congestion on the roads even with the reduced traffic demand and a good spatial and temporal coverage provided by the railway network. Positive values for $\beta_{\text{car availability}}$ and $\beta_{\text{high income}}$ indicate that people owning a car or having higher incomes prefer to use car as their mode of choice as expected. A positive $\beta_{\text{old age}}$ in rail utility function indicates that old people prefer to use railway over car which is as expected. A negative value of $\beta_{\text{work in suburbs}}$ in the rail utility function indicates that people working in suburbs prefer to use car over the rail. This is also plausible due to the fact that they mostly commute to industrial areas out of the city sparsely populated and with lesser railway coverage than the central Tokyo. A negative value of the β_{young} in railway utility function indicates that young people prefer to use car over the railway. The definition of young in this case is people less than 30 years of age, which are mostly either students or company workers just starting their careers. This trend can be explained as a counter to the old people's preference for rail.

4.2 Nested logit models

The MNL models estimated in the previous section, assume no correlation among the alternatives but some of the choices especially in case of departure time choice may be intrinsically correlated. To capture the effect of these correlations, different nested logit structures as depicted in the Figure 2 to Figure 5 are estimated using the same data and the utility specifications including the level of service and personal attributes.

The results of these four nested logit models are reported in Table 3. The results were estimated using the MNL parameters as initial values. All the parameters are significant in all the four models at more than 95% significance level except the $\beta_{\text{old age}}$ in nested model with 4 nests where it is significant at 94% level. Results indicate that all the four nesting structures are significant. Statistical tests indicate that nesting parameter in all the four models is significantly different from null and unit hypothesis values. Likelihood ratio tests comparing all the nesting models to corresponding MNL model are satisfied at more than 99th percentile of a χ^2 random variable with one degree of freedom. The value of travel time shows some differences from the MNL model, for example in case of 2 nests it is higher while in all the other nests it is lower than the MNL model. Values for early arrival and late arrival penalties are also found to be resilient to the changes in correlation structure.

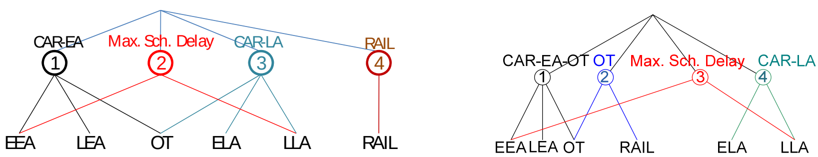
TABLE 3: Estimation results of NL models of Figures 2, 3, 4, and 5

Coefficients	2 Nest NL model (Figure 2)		3 Nest NL model (Figure 3)		3 Nest NL model (Figure 4)		4 Nest NL model (Figure 5)	
ASC_{Rail}	2.7988	(5.04)	5.1608	(3.66)	2.9832	(5.24)	5.0572	(3.25)
$\beta_{travel\ time}$	-0.0414	(-3.78)	-0.0785	(-3.54)	-0.0495	(-3.48)	-0.0767	(-3.15)
β_{cost}	-0.0010	(-3.30)	-0.0025	(-3.53)	-0.0014	(-3.22)	-0.0020	(-3.07)
$\beta_{early\ arrival}$	-0.0231	(-6.51)	-0.0463	(-4.68)	-0.0309	(-4.63)	-0.0458	(-3.99)
$\beta_{late\ arrival}$	-0.1026	(-7.15)	-0.2616	(-4.04)	-0.1587	(-4.31)	-0.2455	(-3.60)
$\beta_{car\ availability}$	1.4848	(4.54)	2.8604	(3.43)	1.5469	(4.68)	2.7300	(3.08)
$\beta_{old\ age}$	2.0729	(2.36)	3.9534	(2.00)	2.2378	(2.20)	3.7227	(1.92)*
$\beta_{high\ income}$	1.7708	(3.41)	3.4333	(2.94)	2.0178	(3.44)	3.2418	(2.75)
β_{young}	-0.8750	(-2.65)	-1.6697	(-2.29)	-1.0335	(-2.47)	-1.6117	(-2.18)
$\beta_{work\ in\ suburbs}$	-1.6012	(-4.74)	-3.1012	(-3.58)	-1.6217	(-5.54)	-2.9553	(-3.28)
$\mu(0)$	0.6275	(5.67)	0.3254	(3.98)	0.5486	(4.67)	0.3461	(3.49)
(1)		(-3.37)		(-8.26)		(-3.84)		(-6.60)
No. of observations	1324		1324		1324		1324	
No. of parameters	11		11		11		11	
Null-log likelihood	-1454.6		-1454.6		-1454.6		-1454.6	
Final-log likelihood	-978.9		-969.8		-977.4		-974.6	
Rho-Squared	0.327		0.333		0.328		0.330	
Rho-Squared bar	0.319		0.326		0.321		0.322	
VTTS(yen/min)	42.2		31.1		35.8		37.7	
VEAP(yen/min)	23.5		18.4		22.4		22.5	
VLAP(yen/min)	104.5		103.8		114.9		120.6	

Nested logit model shown by the correlation structure in Figure 3 performs better than the other nesting structures. The fact that other nesting structures are also significant indicates the presence of cross-nesting which is investigated in next section.

4.3 Cross-nested logit models

As reported in Section 4.2, four different correlation structures are found to be significant indicating the possibility of the cross-nesting among different alternative sets. To test this hypothesis, we developed different cross-nesting structures by combining the nesting structures shown in Figures 2-5 and two of these have been found to be significantly better than the nested models (see Figure 6).



(A) (B)

FIGURE 6: Two cross-nesting structures

Cross-nesting structure (A) shows that departure time alternatives using car are not only grouped together based on their order but also by schedule delay associated with them On-time arrival alternative also belongs to two nests i.e. early and late arrival nests indicating its proximity to both. Cross-nesting structure (B) also indicates that departure time alternatives using car are not only grouped together based on their order but also by schedule delay associated with them. For example, nest 1 shows the alternatives of early arrival and on-time arrival using car as grouped together and nest 4 shows that alternatives having late arrival time are grouped together but at the same time Earliest Early Arrival (EEA) and Latest Late Arrival (LLA) are also found to be nesting together in nest 3 indicating that departure time alternatives are not only correlated due to their proximity to each other but also due to the schedule delay associated with them. This nesting by schedule delay hypothesis is further strengthened by the observation that on-time arrival by car which shares a nest with early arrival using car options also share nest with the on-time arrival using rail.

Table 4 shows the estimation results for the two cross-nesting structures shown in Figure 6. All the attribute parameters, nesting parameters (μ's) as well as cross-nest share parameters (α's) of alternatives are found to be significant. The log-likelihood values show significant improvement over the nested logit models shown in Section 4.2 with one and three extra parameters for cross-nesting structure (A) and (B).

TABLE 4: Estimation results of cross-nested logit models

Coefficients	Cross-nesting structure (A)		Cross-nesting structure (B)	
ASC_{Rail}	10.2	(3.44)	1.037	(3.15)
$\beta_{\text{travel time}}$	-0.172	(-2.76)	-0.1692	(-2.8)
β_{cost}	-0.00541	(-2.73)	-0.0046	(-3.28)
$\beta_{\text{early arrival}}$	-0.0853	(-3.62)	-0.0830	(-3.82)
$\beta_{\text{late arrival}}$	-0.516	(-3.52)	-0.516	(-3.57)
$\beta_{\text{car availability}}$	5.46	(3.29)	5.5140	(3.10)
$\beta_{\text{old age}}$	7.61	(2.12)	7.7750	(2.13)
$\beta_{\text{high income}}$	6.38	(2.66)	6.7083	(2.68)
β_{young}	-3.2	(-2.26)	-3.0507	(-2.12)
$\beta_{\text{work in suburbs}}$	-5.92	(-3.35)	-5.9807	(-3.22)
$\mu(0)$	0.168	(3.57)	0.166	(3.58)
(1)		(-17.6)		(-17.95)
$\alpha_{EEA,1}(1)$	0.734	(-2.86)	0.762	(-2.6)
$\alpha_{EEA,2}(1)$	0.266	(-7.91)	--	--
$\alpha_{EEA,3}(1)$	--	--	0.238	(-8.3)
$\alpha_{OT,1}(1)$	0.505	(-6.28)	0.745	(-3.8)
$\alpha_{OT,2}(1)$	--	--	0.255	(-11.1)
$\alpha_{OT,3}(1)$	0.495	(-6.41)	--	--
$\alpha_{LLA,2}(1)$	--	--	--	--
$\alpha_{LLA,3}(1)$	0.556	(-2.94)	0.464	(-3.5)
$\alpha_{LLA,4}(1)$	0.444	(-3.67)	0.536	(-3.04)
No. of observations	1324		1324	
No. of parameters	17		14	
Null-log likelihood	-1454.56		-1454.56	
Final-log likelihood	-958.7		-958.56	
Rho-Squared	0.341		0.341	
Rho-Squared bar	0.329		0.329	
VTTS(yen/min)	31.8		36.8	
VEAP(yen/min)	15.8		18.0	
VLAP(yen/min)	95.4		112.2	

Values of travel time savings for cross-nesting structure (A) is 28 yen/min while the value of early arrival penalty and late arrival penalty are 19 and 98 yen/min respectively. Nesting structure (B) reflect a value of travel time savings of around 37 yen/min and values of late and early arrival penalties of around 18 and 112 yen/min. Cross-nesting parameters (α's) are all significantly different from zero and one, hence indicating that they are not dominantly contained in a single nest. It also confirms the fact that different correlation structures co-exist among different alternatives as indicated separately by different nesting structures being significant at the same time in previous section.

In order to evaluate the relative improvement in model performance due to an additional nest for maximum schedule delay, cross-nested model shown in Figure 6 (A) is compared with another model without the cross-nesting structure. These model structures are shown in Figure 7. Estimation results for these models are provided in Table 5.

TABLE 5: Estimation results for comparison of cross-nested logit models

Coefficients	Cross-nesting structure Figure 7 (A)		Cross-nesting structure – relaxed Figure 7 (B)	
ASC_{Rail}	10.2	(3.44)	3.14	(4.91)
$\beta_{\text{travel time}}$	-0.172	(-2.76)	-0.0588	(-3.7)
β_{cost}	-0.00541	(-2.73)	-0.00198	(-3.48)
$\beta_{\text{early arrival}}$	-0.0853	(-3.62)	-0.0193	(-4.55)
$\beta_{\text{late arrival}}$	-0.516	(-3.52)	-0.139	(-5.16)
$\beta_{\text{car availability}}$	5.46	(3.29)	1.61	(4.31)
$\beta_{\text{old age}}$	7.61	(2.12)	2.20	(2.37)
$\beta_{\text{high income}}$	6.38	(2.66)	1.91	(3.19)
β_{young}	-3.2	(-2.26)	-0.944	(-2.55)
$\beta_{\text{work in suburbs}}$	-5.92	(-3.35)	-1.75	(-4.59)
$\mu(0)$	0.168	(3.57)	0.58	(5.54)
(1)		(-17.6)		(-4.01)
$\alpha_{EEA,1}(1)$	0.734	(-2.86)	--	--
$\alpha_{EEA,2}(1)$	0.266	(-7.91)	--	--
$\alpha_{EEA,3}(1)$	--	--	--	--
$\alpha_{OT,1}(1)$	0.505	(-6.28)	1.0	--
$\alpha_{OT,2}(1)$	--	--	0.0	--
$\alpha_{OT,3}(1)$	0.495	(-6.41)	--	--
$\alpha_{LLA,2}(1)$	--	--	--	--
$\alpha_{LLA,3}(1)$	0.556	(-2.94)	--	--
$\alpha_{LLA,4}(1)$	0.444	(-3.67)	--	--
No. of observations	1324		1324	
No. of parameters	17		13	
Null-log likelihood	-1454.56		-1454.56	
Final-log likelihood	-958.7		-968.47	
Rho-Squared	0.341		0.334	
Rho-Squared bar	0.329		0.325	
VTTS(yen/min)	31.8		29.7	
VEAP(yen/min)	15.8		9.8	
VLAP(yen/min)	95.4		70.2	

Results in Table 5 compare the two models with and without the maximum schedule delay nest and the comparative results indicate that the model with the maximum schedule delay nest is better than the model without maximum schedule delay nest. This result is also confirmed using the likelihood ratio test. Similar analysis was also conducted for smaller schedule delays. It has been found that addition of another nest for smaller schedule delays does not improve significance of the model and the small early and late schedule delays are not correlated at all. This can be explained owing to different perceptions attached to small early or late schedule delays. For example, being early by 15 minutes may be less onerous compared to being late for 15 minutes and people prefer to arrive a little earlier than being late indicating different perceptions and hence no correlation for smaller early and late schedule delays. On the other hand, the longer schedule delays are equally undesirable no matter whether early or late and are hence correlated.

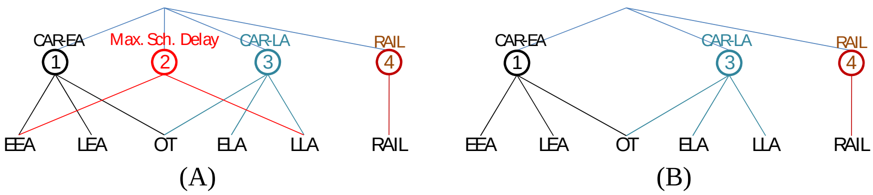


FIGURE 7: Cross-nesting structures with and without maximum schedule delay nest

4.4 Random coefficient multinomial logit models

As discussed before, it is important to account for the random taste variations across the individuals. We tried to explore the attributes that are perceived and treated differently among the population of the commuters. The variables indicating late arrival and work in suburbs have significant random coefficients. It is assumed that these random coefficients are normally distributed.

Results of random coefficient estimation for different random coefficient specifications are shown in Table 6. First mixed MNL model is built using work in suburbs as random variable. All the parameters are significant at more than the 95% significant level but the overall improvement in the model fitness is not very high (2(L(MMNL)-L(MNL)) = 3.0 > 2.706, 90th percentile of a χ^2 random variable with one degree of freedom). The value of travel time savings as well as value of early arrival and late arrival penalties remains same as the MNL model.

Second model is estimated using work in suburbs as well as the late arrival as random variables. All the parameters are significant at the 95% significance level. The improvement in log-likelihood is about 25 units over the MNL model indicating that this mixed MNL is better than MNL model at 99% significance level indicating important gains in model performance obtained by using the random coefficient model. The normally distributed work in suburbs parameter has a distribution of N(-0.994,2.14) indicating that about 32% of the commuters who work in suburbs have positive utility for rail. This effect was not captured by using the fixed parameter MNL model. On the other hand, it can be noted that use of normal distribution for the late arrival penalty results in a parameter distribution of N(-0.3023,0.17), which indicates that about 4% of the commuters get positive utility from being late which is counter-intuitive. Although this number is not very high but it would be better if a log-normal distribution is tried which is constrained to remain in a single sign domain. Another interesting result depicted by this model is a very high value for the late arrival penalty which is more than double the value in MNL model with a very broad distribution. The Value of late arrival penalty has a distribution of N(252, 148). Higher variance in late arrival distribution can be explained by differences of personal preferences as the value of late arrival penalty may be very high for an office worker or executive while can be low for a student or people with flexible work schedules. This effect can only be modelled using the random coefficient models.

Third model described in Table 6 shows the results for mixed logit estimation; early and late arrival and work in suburb parameters are normally distributed. Results indicate a decrease in the significance level for some of the parameters below 95% level. Also, it is clear that the early arrival penalty has a very narrow though significant distribution *N*(-0.0217,0.0011) indicating that a point estimate of the parameter is enough to represent it. Hence, model including the work in suburbs as well as late arrival penalty as random coefficients is retained for the further investigations.

TABLE 6: Estimation results of mixed MNL model of Figure 1

Coefficients	Work in suburbs as random variable		Work in suburbs + late arrival as random variable		Late arrival + early arrival as random variable	
ASC_{Rail}	1.8588	(8.59)	1.6035	(6.24)	1.6090	(7.83)
$\beta_{\text{travel time}}$	-0.0337	(-3.85)	-0.0380	(-3.87)	-0.0372	(-4.35)
β_{cost}	-0.0009	(-3.32)	-0.0012	(-3.73)	-0.0009	(-3.25)
$\beta_{\text{early arrival}}$	-0.0198	(-6.94)	-0.0273	(-8.30)	-0.0217	(-8.63)
$\sigma_{\text{early arrival}}$	--	--	--	--	0.0011	(2.04)
$\beta_{\text{late arrival}}$	-0.0936	(-7.65)	-0.3023	(-5.07)	-0.2545	(-5.10)
$\sigma_{\text{late arrival}}$	--	--	-0.17	(2.99)	-0.1504	(-4.65)
$\beta_{\text{car availability}}$	1.0421	(6.34)	1.1042	(5.86)	0.8612	(5.96)
$\beta_{\text{old age}}$	1.5960	(2.52)	2.0469	(2.99)	0.5516	(1.53)*
$\beta_{\text{high income}}$	1.1968	(4.16)	1.1761	(3.80)	0.5058	(1.68)*
β_{young}	-0.5784	(-2.60)	-0.5792	(-2.25)	-0.2860	(-1.39)*
$\beta_{\text{work in suburbs}}$	-0.9832	(-6.13)	-0.9942	(-5.16)	-1.013	(-7.50)
$\sigma_{\text{work in suburbs}}$	1.4377	(2.26)	2.1410	(2.60)	--	--
No. of observations	1324		1324		1324	
No. of parameters	11		12		12	
Null-log likelihood	-1454.6		-1454.6		-1454.6	
Final-log likelihood	-981.3		-958.84		-968.32	
Rho-Squared	0.325		0.341		0.334	
Rho-Squared bar	0.320		0.333		0.326	
VTTS(yen/min)	37.4		31.7		41.3	
VEAP Mean (yen/min)	22		22.8		24.1	
VEAP Variance (yen/min)	--		--		1.2	
VLAP Mean (yen/min)	104		251.9		282.8	
VLAP Variance (yen/min)	--		147.7		167.1	

4.5 Random coefficient nested logit models

Nested logit models described in Section 4.2 provide an improvement over the simple MNL model by capturing the correlation among the alternatives while the mixed logit models described in Section 4.3 improve upon the MNL model by accounting for the random taste heterogeneity. In this section, we combine these two types of models to jointly account for the correlation among the alternatives as well as the random taste variations among the population of the commuters. Same nesting structures as described in Section 3.4 and used in Section 4.2 are employed here and corresponding mixed nested logit model are estimated.

Table 7 details the results of the estimations for the four nesting structures. All the parameters are statistically significant at 95% confidence level except the variance for the work in suburbs random variable which loses significance at any suitable level for two of the nesting structures.

TABLE 7: Estimation results of mixed NL models of Figures 2, 3, 4, and 5

Coefficients	2 nest NL model (Figure 2)		3 nest NL model (Figure 3)		3 nest NL model (Figure 4)		4 nest NL model (Figure 5)	
ASC_{Rail}	5.1259	(3.45)	3.2958	(4.65)	4.1796	(3.41)	3.8949	(3.39)
$\beta_{travel\ time}$	-0.0810	(-3.73)	-0.0701	(-3.50)	-0.0804	(-3.61)	-0.0783	(-3.25)
β_{cost}	-0.0025	(-3.54)	-0.0023	(-3.84)	-0.0027	(-3.65)	-0.0023	(-3.42)
$\beta_{early\ arrival}$	-0.0531	(-5.53)	-0.0525	(-5.65)	-0.0525	(-5.27)	-0.0528	(-4.74)
$\beta_{late\ arrival}$	-0.5833	(-4.02)	-0.6575	(-3.82)	-0.7150	(-3.23)	-0.6589	(-3.42)
$\sigma_{late\ arrival}$	-0.3609	(-3.66)	-0.3884	(-3.62)	-0.4235	(-3.05)	-0.3894	(-3.25)
$\beta_{car\ availability}$	3.1643	(3.45)	2.1580	(4.50)	2.7010	(3.66)	2.5783	(3.41)
$\beta_{old\ age}$	4.6176	(2.28)	4.3912	(2.64)	4.1469	(2.47)	4.3479	(2.47)
$\beta_{high\ income}$	3.6260	(2.86)	2.4600	(3.64)	3.0031	(2.95)	2.7488	(2.89)
β_{young}	-1.8084	(-2.25)	-1.3416	(-2.24)	-1.5085	(-2.16)	-1.3895	(-2.04)
$\beta_{work\ in\ suburbs}$	-3.3823	(-3.57)	-2.0117	(-4.17)	-2.7152	(-3.15)	-2.3780	(-3.35)
$\sigma_{work\ in\ suburbs}$	0.3484	(0.21)*	4.6790	(2.87)	2.5022	(0.7)*	3.9974	(1.96)
$\mu(0)$	0.2976	(3.88)	0.4394	(5.37)	0.3711	(3.77)	0.4157	(3.99)
(1)		(-9.16)		(-6.85)		(-6.39)		(-5.61)
No. of observations	1324		1324		1324		1324	
No. of parameters	13		13		13		13	
Null-log likelihood	-1454.6		-1454.6		-1454.6		-1454.6	
Final-log likelihood	-945.4		-948.4		-948.3		-951.8	
Rho-Squared	0.350		0.348		0.348		0.346	
Rho-Squared bar	0.341		0.339		0.339		0.337	
VTTS(yen/min)	32.4		30.5		29.8		34	
VEAP(yen/min)	21.2		22.8		19.5		23	
VLAP(yen/min)	233.3		286		265		287	
VLAP Variance (yen/min)	144.4		169		157		169	

Comparison of the estimated model results with the corresponding mixed MNL model indicates an improvement in the overall model fitness at a significance level of over 99.5% with a single degree of freedom indicating the gains in performance of the model. Similar observations can be made by comparing the mixed nested logit models with corresponding nested logit models in Section 4.2. The improvement in log-likelihood over the nested logit models of the Section 4.2 with 2 degrees of freedom is statistically significant at 99.5% level.

The value of travel time savings is reduced in comparison to MNL and NL models and is around 30 to 35 yen per min in this case while the value of early arrival penalty is consistent at about 20 to 23 yen/min. The value of late arrival penalty show a distribution with a mean of around 250 to 300 yen and corresponding variance of around 140-170 yen which is consistent with the results obtained for the mixed MNL model in the previous section.

Mixed NL model corresponding to nesting structure shown in Figure 2 has better goodness-of-fit compared to other models but one of the parameters is insignificant. Mixed NL models corresponding to nesting structures shown in Figures 3 and 4 show similar goodness-of-fit measure. In case of NL models, nesting structure corresponding to Figure 3 had better goodness-of-fit than other models and in case of mixed NL models, this nesting structure again has the best goodness-of-fit with all parameters being significant. One of the trends observed in all the above proposed models is a gradual decrease in the significance levels of the parameters though they are still significant at 95% confidence level. This is quite expected as each subsequent modelling structure introduced above, decomposes the error term further than the previous models.

4.6 Error component logit models

Mixed logit models can also be explained from another interpretational viewpoint and that is to model the correlation structures among the alternatives. Error component logit models can virtually approximate any other modelling structure for discrete choices provided appropriate specifications are used. This section describes four error component logit models corresponding to four nested logit structures depicted in Section 3.4 and used in Sections 4.2 and 4.4. Random error terms are introduced across alternatives which remain constant for a given set of alternatives hence, indicating the correlation among them.

Table 8 shows the results of the estimations for the error component logit models corresponding to different nested logit correlation structures. Model results indicate that for three out of four correlation structures (namely three and four nest models) error component logit outperforms the corresponding nested logit models while for one (2 nest model) it is statistically same at a higher significance level. This shows that error component logit models can better capture the correlation structures as compared to the nested logit models.

TABLE 8: Estimation results of error component logit models

Coefficients	Corresponding to 2 nest NL model (Figure 2)		Corresponding to 3 nest NL model (Figure 3)		Corresponding to 3 nest NL model (Figure 4)		Corresponding to 4 nest NL model (Figure 5)	
ASC_{Rail}	2.7484	(4.84)	7.5034	(2.90)	3.1556	(7.98)	7.6417	(2.74)
$\beta_{travel\ time}$	-0.0409	(-3.72)	-0.0794	(-3.24)	-0.0430	(-3.51)	-0.0824	(-2.98)
β_{cost}	-0.0010	(-3.28)	-0.0027	(-3.47)	-0.0015	(-3.54)	-0.0024	(-3.24)
$\beta_{early\ arrival}$	-0.0230	(-6.38)	-0.0510	(-4.58)	-0.0388	(-5.44)	-0.0570	(-4.51)
$\beta_{late\ arrival}$	-0.1023	(-7.10)	-0.1253	(-3.28)	-0.0593	(-3.81)	-0.1413	(-2.92)
$\beta_{car\ availability}$	1.4354	(4.43)	3.4943	(2.81)	1.4040	(5.87)	3.5820	(2.56)
$\beta_{old\ age}$	1.8382	(2.37)	3.6231	(1.5)*	1.9122	(1.8)*	3.8950	(1.54)
$\beta_{high\ income}$	1.7812	(3.24)	4.5224	(2.69)	1.9833	(4.51)	4.7666	(2.48)
β_{young}	-0.8291	(-2.55)	-2.1808	(-2.12)	-0.9903	(-3.03)	-2.4483	(-2.05)
$\beta_{work\ in\ suburbs}$	-1.5874	(-4.51)	-2.92	(-2.92)	-1.4818	(-7.06)	-3.9019	(-2.62)
ξ_{car}	-0.1145	(-0.2)*						
ξ_{rail}	1.9681	(3.08)	4.4354	(2.66)			4.9559	(2.32)
$\xi_{on\text{-}time}$					0.0134	(0.12)*	-0.0116	(-0.12)*
$\xi_{early\ arrival}$					3.7514	(4.42)	6.4209	(2.85)
$\xi_{late\ arrival}$			0.5276	(0.9)*	-0.0123	(-0.3)*	-0.1413	(1.6)*
$\xi_{early\ arrival/on\text{-}time}$			6.5455	(2.87)				
No. of observations	1324		1324		1324		1324	
No. of parameters	12		13		13		14	
Null-log likelihood	-1454.6		-1454.6		-1454.6		-1454.6	
Final-log likelihood	-979.6		-965.3		-967.7		-961.6	
Rho-Squared	0.326		0.336		0.335		0.339	
Rho-Squared bar	0.318		0.327		0.326		0.329	
VTTS(yen/min)	41		29.4		29		34.3	
VEAP(yen/min)	23		19		26		24	
VLAP(yen/min)	102		46		40		59	

All the random error component terms are assumed as normally distributed with mean zero i.e. $N(0,\sigma)$ where σ is estimated. Results indicate that not all the random error terms are statistically significant. The other parameters are usually significant at 95% confidence level except $\beta_{old\ age}$ whose significance reduces to 85% and 90% confidence levels for two of the models. Value of travel time savings as well as value of early arrival penalty remains almost same as in previous models while the value of late arrival penalty reduces significantly in this model. This may have resulted due to interaction among the late arrival penalty parameter and the error components.

Comparison of these error component logit models with the mixed logit model shows them at par but inferior to the mixed nested logit models. This may be explained by the fact that proposed error component structures just capture the correlation across the alternatives while the mixed nested logit models also account for the random taste heterogeneity in addition to inter-alternative correlation structures.

5. OVERVIEW OF THE ESTIMATION RESULTS

All model estimation results are summarized in Table 9. The table shows the final log-likelihood values and the number of estimated parameters.

TABLE 9: Summary of estimation results

Model	Final log-likelihood	Number of estimated parameters
Null	-1454.6	0
MNL with LOS variables only	-1052.8	5
MNL with both SE and LOS variables	-982.9	10
NL_1 (structure defined in Figure 2)	-978.9	11
NL_2 (structure defined in Figure 3)	-977.4	11
NL_3 (structure defined in Figure 4)	-969.8	11
NL_4 (structure defined in Figure 5)	-974.6	11
Cross nesting structure (A)	-958.7	17
Cross nesting structure (B)	-958.5	14
Mixed MNL (work in suburbs)	-981.3	11
Mixed MNL (work in suburbs + early arrival)	-958.8	12
Mixed MNL (early arrival + late arrival)	-968.3	12
Mixed_NL_1	-945.4	13
Mixed_NL_2	-948.4	13
Mixed_NL_3	-948.3	13
Mixed_NL_4	-951.8	13
Error components_NL_1	-979.6	12
Error components_NL_2	-965.3	13
Error components_NL_3	-967.7	13
Error components_NL_4	-961.6	14

It can be inferred from the results that inclusion of both socio-demographic i.e. personal characteristics and level of service attributes greatly improves the model performance. This result was verified for the simple MNL model. Four different nesting structures were tested to account for correlation among alternatives and all of them were found to be significantly better than the simple MNL structure. Significance of all the four nesting structures indicated the presence of cross-nesting. Different cross-nesting structures were tried and two structures were found significant. All the attribute coefficients, nesting parameters and cross-nest share parameters of the alternatives were found to be significant.

Random coefficient models using few normally distributed parameters were estimated and were found to perform better than corresponding MNL models indicating that they can capture the taste heterogeneity across the users. Only two attributes namely late arrival penalty and work in suburbs were found to have significant random coefficients.

Taste heterogeneity was found to be significant in case of work in suburbs attribute where the distribution of the parameter indicated that about 32% of the commuters will have a positive utility for using trains to work while in the case of the MNL model the parameter estimated for work in suburbs had a negative value indicating the reduction in utility for all the commuters in the population.

Taste heterogeneity also indicated distributed value of late arrival penalty across the individuals, indicating differences between the commuters. This can be explained by the existence of different commuter groups such as office workers or executives for whom it is important to arrive on-time in contrast to a student or a worker with a flexible arrival time indicating a lower value of late arrival penalty.

Mixed nested logit models perform better than the nested logit and mixed MNL models, indicating that they can jointly capture the correlation structures as well as random taste heterogeneity.

Error component logit models were developed corresponding to the correlation structures of nested logit models and mostly perform better than corresponding nested logit models. However, these models did not outperform the mixed nested logit models.

6. CONCLUSIONS

In this paper, results from a comprehensive study of mode and departure time choice problem are presented. The database was formed from a SP study performed among Tokyo commuters. The paper tested both closed-form GEV structures (MNL, NL and CNL) and more advanced mixed logit forms. In this aspect, the main contribution of the paper is to provide a step-by-step comparison of the models, in an attempt to illustrate the relative improvement of the more general model forms.

As expected, nested logit models perform better than the MNL. With respect to the four proposed nesting structures, nesting structure shown in Figure 3, having three nests of rail, car with early/on-time arrival and car with late arrival performs better than other nesting structures but the other nesting structures are also found to be found significant indicating the possibility of different alternatives belonging to more than one nest. Further analysis using CNL models with different nesting structures confirmed it. Reported results show that cross-nesting of the alternatives is significant and cross-nested logit models show significant improvement in model fitness over corresponding nested logit models. It should be stressed that as with the NL models, several possible CNL model structures can be specified, and not all specifications result in statistically significant nesting coefficients.

This paper tested different logit kernel specifications for the mixed logit, namely the MNL and NL models. The mixed MNL performs better than the MNL by accounting for the random taste variations, and similarly the mixed nested logit models perform better than NL models. In addition, the mixed NL model outperforms the mixed MNL model by accounting for both correlation structure as well as random taste variations jointly. Finally, Error component logit corresponding to the correlation structures of the NL models were developed and although they perform better than corresponding NL models, they were not found superior than mixed NL models.

Level of service variables such as travel time, cost and schedule delay as well as socioeconomic characteristics such as age, income, car availability and work locations are found to be significantly affecting the mode and departure time choices of commuters. It has been found that the inclusion of socioeconomic characteristics not only significantly improves the statistical significance of the estimated models but also greatly enhances the explanatory power of these behavioural models. The paper shows that departure time alternatives are correlated by the amount of delay. Results indicate that the late arrival penalties for the commuters are higher in the Tokyo Metropolitan Area. Choice of departure time is found to be sensitive to the schedule constraints as well as congestion levels and costs while the choice of mode is found sensitive to the age, income level, car availability as well as work locations of the commuters.

The models estimated in this paper will be incorporated in a dynamic multi-modal transport simulation model. Given the results obtained in this paper, it has been found to be important to include the behavioural characteristics of the commuters in the multimodal network assignment models. Such behavioural multimodal network assignment models are important to bridge the gap between disaggregate behavioural models and the network assignment models which are mostly based on level of service variables. Although the mixed multinomial logit models have been found to perform slightly better than the advanced closed-form models such as cross-nested logit models but for practical application closed-form models may be advantageous due to relative simplicity and ease to incorporate them in network assignment models. This dynamic multi-modal transport simulation model will be applied to the multi-modal transportation network in the Tokyo Metropolitan Area.

ACKNOWLEDGEMENTS

Authors are thankful to the two anonymous referees who provided insightful comments to help improve this paper.

REFERENCES

Abkowitz, M. (1981) An analysis of the commuter departure time decision. Transportation, 10, 283-297.

Bhat, C. (1998a) Analysis of travel mode and departure time choice for urban shopping trips. Transportation Research Part B, 32, 361-371.

Bhat, C. (1998b) Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice. Transportation Research Part B, 32, 455-466.

Bhat, C. and Steed, J. (2002) A continuous-time model of departure time choice for urban shopping trips. Transportation Research Part B, 36, 207-224.

Bierlaire, M. (2005) An Introduction to BIOGEME (Version 1.4). http://biogeme.epfl.ch/doc/tutorial.pdf.

Boyd, J. and Mellman, J. (1980) The effect of fuel economy standards on the U.S. automotive market: a hedonic demand analysis. Transportation Research Part A, 14, 367-378.

Cardell, N. and Dunbar, F. (1980) Measuring the societal impact of automobile downsizing. Transportation Research Part A, 14, 432-434.

Cascetta, E. and Papola, A. (2003) A joint mode-transit service choice model incorporating the effect of regional transport service timetables. Transportation Research Part B, 37, 595-614.

Chin, A. (1990) Influences on commuter trip departure time decisions in Singapore. Transportation Research Part A, 24, 321-333.

de Jong, G., Daly, A., Pieters, M., Vellay, C., Bradley, M. and Hofman, F. (2003) A model for time of day and mode choice using error components logit. Transportation Research Part E, 39, 245-268.

de Palma, A., Ben-Akiva, M., Lefevre, C. and Litnas, N. (1983) Stochastic equilibrium model of peak period traffic congestion. Transportation Science, 17, 430-453.

Fargier, P. (1981) Effects of the choice of departure time on road traffic congestion. Proceedings of the Eighth International Symposium on Transportation and traffic Theory, Toronto, Canada.

Henderson, J. (1974) Road congestion: a reconsideration of pricing theory. Journal of Urban Economics, 1, 346-365.

Hendrikson, C. and Kocur, G. (1981) Schedule delay and departure time decisions in a deterministic model. Transportation Science, 15, 62-77.

Hendrickson, C. and Plank, E. (1984) The flexibility of departure times for work trips. Transportation Research Part A, 18, 25-36.

Hendrickson, C., Nagin, D. and Plank, E. (1981) Characteristics of travel time and dynamic user equilibrium for travel-to-work. Proceedings of the Eighth International Symposium on Transportation and Traffic Theory, Toronto, Canada.

Hess, S., Bierlaire, M. and Polak, J. (2004) Development and application of a mixed cross-nested logit model. European Transport Conference, Strasbourg.

Hurdle, V. (1981) Equilibrium flows on urban freeways. Transportation Science, 5, 255-293.

Koppelman, F.S. and Wen, C.H. (2000) The paired combinatorial logit model: properties, estimation and application. Transportation Research Part B, 34, 75-89.

Kuwahara, M. (1985) A time-dependent network analysis for highway commute traffic in a single core city. Ph.D Thesis, University of California, Berkeley.

Kuwahara, M. and Newell, G. (1987) Queue evolution on freeways leading to a single core city during the morning peak. Proceedings of the 10th International Symposium on Transportation and Traffic Theory, Boston, pp. 21-40.

Louviere, J., Hensher, D. and Swait, J. (2000) Stated Choice Methods: Analysis and Applications in Marketing, Transportation and Environmental Valuation. Cambridge University Press, Cambridge.

McFadden, D. (1978) Modeling the choice of residential location. In A. Karlqvist et al. (eds.) Spatial Interaction Theory and Residential Location, North-Holland, Amsterdam.

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

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

Polak, J.W. and Jones, P.M. (1994) A tour-based model of journey scheduling under road pricing. 73rd Annual Meeting of the Transportation Research Board, Washington, DC.

Small, K. (1982) The scheduling of consumer activities: work trips. The American Economic Review, 72,, 467-479.

Small, K. (1987) A discrete choice model for ordered alternatives. Econometrica, 55, 409-424.

Shimizu, T. and Yai, T. (1999) An analysis of the effect on peak period toll in Tokyo Metropolitan Expressway. Journal of the Eastern Asia Society for Transportation Studies, 3, 305-318.

Train, K. (2003) Discrete Choice Methods with Simulation. Cambridge University Press, New York.

Vickrey, W. (1969) Congestion theory and transport investment. The American Economic Review, 59, 251-261.

Vovsha, P. (1997) Application of a cross-nested logit model to mode choice in Tel Aviv, Israel, Metropolitan Area. Transportation Research Record, 1607, 6-15.

Vovsha, P. and Bekhor, S. (1998) The link-nested logit model of route choice: overcoming the route overlapping problem. Transportation Research Record, 1645, 133-142.

Walker, J. (2001) Extended discrete choice models: integrated framework, flexible error structures and latent variable. Ph.D. Thesis, Massachusetts Institute of Technology.

Wang, J. (1996) Timing utility of daily activities and its impact on travel. Transportation Research Part A, 30, 189-206.

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