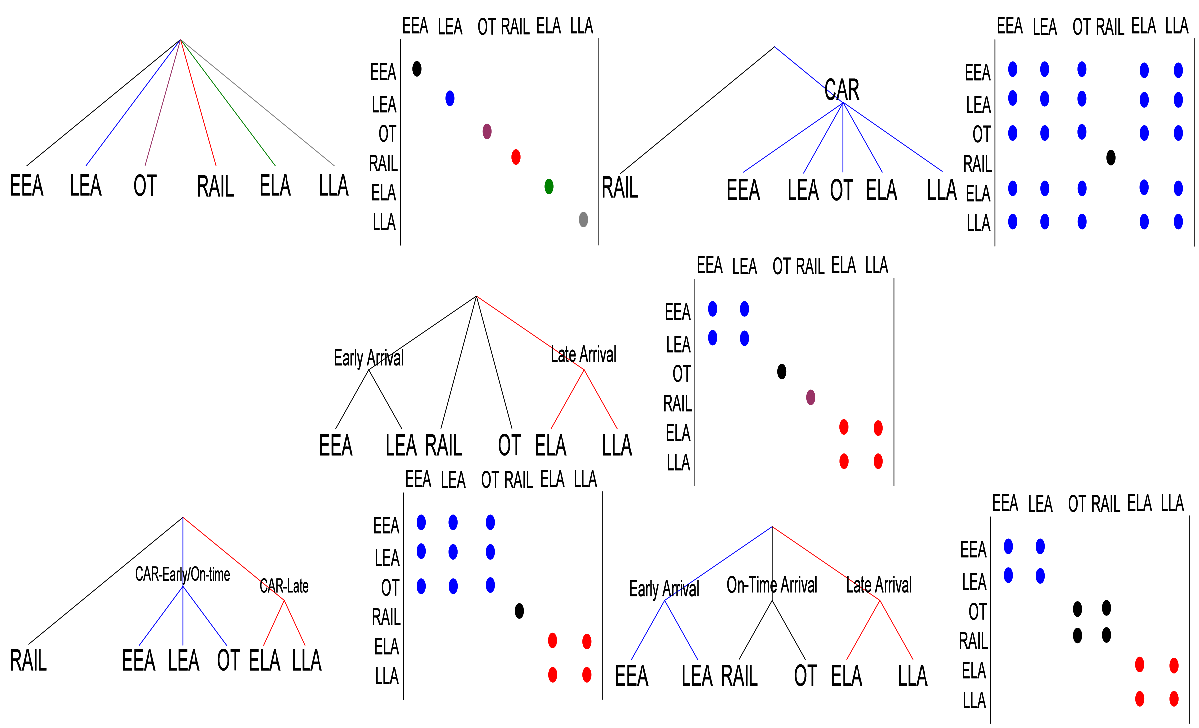

Discrete Choice Modeling of Combined Mode and Departure Time

Shamas ul Islam Bajwa, University of Tokyo
Shlomo Bekhor, Technion – Israel Institute of Technology
Masao Kuwahara, University of Tokyo
Edward Chung, University of Tokyo

Conference paper
Session XXX

The Expanding Sphere of Travel Behaviour Research

11th International Conference on Travel Behaviour Research
Kyoto, 16-20, August 2006

Discrete Choice Modeling of Combined Mode and Departure Time

Shamas ul Islam Bajwa
Kuwahara Lab, Institute of Industrial Science
University of Tokyo
Japan
Phone: 081 – 3 – 5452 - 6419
Fax: 081 – 3 – 5452 - 6420
E-mail: shamas@iis.u-tokyo.ac.jp

Shlomo Bekhor
Transportation Research Institute
Technion – Israel Institute of Technology
Israel

Masao Kuwahara
Institute of Industrial Science
University of Tokyo
Japan

Edward Chung
Institute of Industrial Science
University of Tokyo
Japan

Abstract

Typical daily decision-making process of individuals regarding use of transport system involves mainly three types of decisions: mode choice, departure time and route choice. This paper focuses on the first two of these decision processes, and details the development of a combined departure time and mode choice model. Data from a stated choice survey of morning commuters conducted in Tokyo, Japan is used for this study. Two different modes i.e. car and rail are modeled in this study. Given the high reliability of the rail system, it is assumed that the rail choice is associated to a reference departure time, while the departure time choice for car users varies relative to the rail time. Different specifications of the model structure including different sets of explanatory variables as well as different model structures to capture the correlation among alternatives as well as taste variations among the commuters are explored. Correlation among different alternatives is confirmed by trying different nesting structures as well as error component formulations. A few trials with random coefficient logit specification also confirm the presence of the random taste heterogeneity across different commuters. To jointly account for the random taste heterogeneity as well as the correlation among different alternatives mixed nested logit models are estimated. Results indicate that accounting for the random taste heterogeneity as well as inter-alternative correlation improves the model performance though some of the coefficients become less significant.

Keywords

Departure Time Choice, Mode Choice, Mixed Logit, Schedule Delay, Value of Time, Nested Logit, Cross-nested Logit, Error Component Logit, Value of Early Arrival Penalty, Value of Late Arrival Penalty

Preferred Citation

Bajwa, Shamas ul Islam, Shlomo Bekhor, Masao Kuwahara, and Edward Chung, (2006) Discrete Choice Modeling of Combined Mode and Departure Time, paper presented at the 11th International Conference on Travel Behaviour Research, Kyoto, August 2006.

1. Introduction

Advent of the advanced transport control and communication technologies has made it possible to implement time-varying demand management policies such as time-varying road pricing and demand responsive traffic control such as ramp-metering. To assess the impact of these policies, it is necessary to develop the behavioural models accounting for these effects.

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.

Usually, all commuters have preferred arrival times at their destinations due to constraints of work starting times. These arrival times are generally concentrated in a narrow band. 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 and/or cost) or arriving on time but spending more time in congestion and/or higher costs. Hence, the parameter of interest is not only the value of travel time savings, as is usually the case in traditional models, but also the value of early schedule delay as well as late schedule delay to establish the trade-off among the different alternatives available to users.

In this paper, our main focus is the departure time choice behavior of road traffic network users. However, it is also known that due to extensive availability and blanket coverage of the alternative public transport modes in Tokyo area, mode choice always stands as a viable option. Need for such mode choice behavior models is further highlighted if we want to test some travel demand management strategies such as road pricing which may be prohibitively expensive for at least some section of the travelers. Hence, to provide a viable alternative to travel, mode choice is also considered. Mode choice can also be considered as a stand-in for elastic demand in the traffic networks as assuming a constant travel time between origin and destination as well as a fixed cost and schedules can keep its utility constant, giving it the role of a null option as is the case for the elastic demand models.

Important considerations for behaviour modelling include appropriate specification of the utility functions associated with each alternative. Users' choices depend not only on their socio-demographic characteristics such as age, gender, income, residence and work locations, but also on the level of service attributes of the network which they use for commuting. An appropriate utility function specification should include a mix of these characteristics 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 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 over the population instead of a fixed single value. In addition, the 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, we aim to both identify appropriate utility specifications and capture the random taste heterogeneity across the commuters by using the mixed logit class of models. In addition, we investigate correlation among alternatives by specifying different nesting structures as well as error components.

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 used. 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.

2. Literature Survey

Almost all the existing models about departure time choice are based on the tradeoff principle proposed by Vickrey (1969) between early, late or on-time arrival. The author proposed a pioneering model for deterministic departure time choice for morning commute, which assumes that *N* identical travelers have the same PAT (Preferred Arrival Time) and use the same route. 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 late and arriving later than PAT. The model formulation represents time as a continuous variable.

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 pattern, 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. The general trade-off principle proposed in these models is similar and is also employed for discrete choice models.

Small (1982) used the data collected from the car commuters in the San Francisco bay area to model the arrival time choice using multinomial logit (MNL) model. A number of socioeconomic and level of service variables, such as family status, occupation, mode of transport, and work hour flexibility, were used. Hendrickson and Plank (1984) also proposed MNL structure for combined mode and departure time choice model. Abkowitz (1981) also used the MNL model on the same data set as used by Small (1982) including additional socio-demographic variables (income and age) and transit mode use as determinants of commuters' departure time choice behavior. Chin (1990) also used MNL model for the departure time choice using the data collected in Singapore of morning commuters. He found that departure time choices were influenced by the travel time as well as occupation and income. Shimizu and Yai (1999) carried out a survey to gauge the reaction of commuters to a variable peak period toll in Tokyo Metropolitan area. They used Structural Equations to model the behavior of the users. They afterwards modeled the departure time choice of the users 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. Usually smaller the departure time interval becomes, more difficult it is for the decision-makers to distinguish between the adjacent time intervals resulting in a higher correlation indicating a problematic model structure. 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 some natural ordering which is time-of-day in this case. 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) form, which recognizes the natural temporal ordering of the departure time alternatives, for departure time choice. The proposed MNL-OGEV model was applied to data obtained from the 1990 San Francisco Bay area travel survey 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.

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 by allowing for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors over time. Unlike probit, it is not restricted to normal distributions. Its derivation is straightforward, and simulation of its choice probabilities is computationally simple (Train, 2003). The first application of mixed logit was apparently the automobile demand models created jointly by Boyd and Mellman (1980) and Cardell and Dunbar (1980) who applied mixed logit for the automobile demand models.

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 highlight the need to capture unobserved attributes along both the mode and departure time dimensions, both for improved data fit as well as for more realistic policy 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 that 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 choice (Wang 1996, Bhat and Steed 2002).

In most researches, the multinomial logit model 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 using a SP data from Switzerland. They propose such modeling structures to capture the effect of the random taste heterogeneity as well as the inter-alternative correlation. According to them, use of mixed GEV models results in important gains in the performance over the use of basic models.

3. Methodology

Discrete choice models are used to replicate the choice 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 tested in this paper. Data describing users' behaviour is presented next. The choice set definition is described next. The last subsection presents the specification of the utility functions and correlation structure.

3.1 General Model Structure

The usual form of the utility function is,

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

where,

U_{ij}: the utility of the individual j for alternative $i \in C_j$,

C_j: choice set available to decision-maker j

V_{ij}: the deterministic part of the utility of the alternative i for individual j,

ε_{ij}: the random component of the alternative i for individual j.

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

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.

3.2 3.1.1 Closed-form GEV Models

McFadden (1978) derived the Generalized Extreme Value (GEV) model from the random utility model. This general model consists of a large family of models that include the Multinomial Logit and the Nested Logit models as special cases. The probability of choosing alternative i within C_j is,

$$P(i \mid C_j) = \frac{e^{V_{ij}} \dfrac{\partial G(e^{V_{1j}}, \ldots\ldots.., e^{V_{Ij}})}{\partial e^{V_{ij}}}}{\mu G(e^{V_{1j}}, \ldots\ldots.., e^{V_{Ij}})} \tag{3}$$

where,

I_j: Number of alternatives in C_j

μ: scale parameter

G: A non-negative differentiable function with following properties(Ben-Akiva and Bierlaire, 1999):

1. G is homogenous of degree $\mu > 0$.
2. $\lim_{x_i \to \infty} G(x_1, \ldots\ldots, x_i, \ldots\ldots x_{Ij}) = \infty, \forall i = 1, \ldots\ldots, I_j$
3. the kth partial derivative with respect to k distinct x_i is non-negative if k is odd, and non-positive if k is even, that is, for any distinct $i_1, \ldots i_k \in \{1, \ldots I_j\}$ we have

$$(-1)^k \frac{\partial^k G(x)}{\partial x_{i_1} \ldots \partial x_{i_k}} \leq 0$$

MNL model can be derived using a simple generating function G, given as,

$$G(x) = \sum_{i=1}^{I_j} x_i^{\mu} \tag{4}$$

NL model with M nests can be derived using the following generating function,

$$G(x) = \sum_{m=1}^{M} \left(\sum_{i \in C_{mi}} x_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \tag{5}$$

where,

C_{mi}: set of alternatives contained in nest m

μ_m: structural parameter of nest m, where $0<\mu/\mu_m<1$

The generating function for the nested logit model can be extended to allow for a cross-nested model, in which one alternative can belong to different nests,

$$G(x) = \sum_{m=1}^{M} \left(\sum_{i \in C_{mi}} (\alpha_{im} x_i)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \tag{6}$$

where,

α_{im}: allocation parameter for alternative i to nest m

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, allowing for 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 and Papola 2003).

3.3 Mixed GEV Models

In a mixed GEV model, the term V_{ij} in equation (1) also becomes random. Hence, the probability of choosing alternative i within C_j is,

$$P(i \mid C_j) = \int P(i \mid C_j, \beta) f(\beta) d\beta \tag{7}$$

where,

$P(i \mid C_j, \beta)$: probability of choosing mode i out of C_j at a given parameter β

$f(\beta)$: density function

Based on above definition, two conceptually different models can be developed; the random coefficient logit (RCL) model, and the error component logit (ECL) model.

In the RCL 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.

$$P(i \mid C_j) = \int_{\beta} P(i \mid C_j, \beta) f(\beta, \theta) d\beta \tag{8}$$

where,
θ: vector of parameters of the distribution of the elements contained in the vector β, for example, mean and standard deviation

In the ECL 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,

$$V_{ij} = f(\beta, x_{ij}) + \xi \tag{9}$$

where,
ξ : 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).

In this case, the random utility contains two error terms,

$$U_{ij} = V_{ij} + \varepsilon_{ij} + \xi \qquad (10)$$

ECL models can capture the correlation among different alternatives by allowing some alternatives to share the same error components.

Though traditionally the mixed GEV models have been Mixed MNL (MMNL) i.e. the generating function used mostly is as in equation (4) but in actual any perceived generating function satisfying the properties described above can be used.

3.4 Survey Design and Data Collection

A stated preference survey was conducted in Tokyo Metropolitan Area to elicit the responses of the commuters 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 is modeled in 15 minute intervals and this option is only available if commuters choose car as their mode. Car commuters can make a tradeoff 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 are also introduced. This allows us to measure the tradeoff of the monetary costs and travel time 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 to avoid severe congestion on 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 consists of the personal characteristics such as age, gender, income, work location (post code) and information about a typical day morning commute. In second part, stated preference scenarios were presented and users were asked to choose departure time as well as mode. Number of mode choices available to the people was two i.e. rail and car while the departure time choice for the rail users was only one while different departure time choice for the car users were presented. 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 owing to ignoring second and higher order effects i.e. interactions among two or more than two attributes but it has been shown that more than 80% of the information is explained by main effects while the 15% is held by the first order effects and remaining 5% is held by the second and higher order effects (Louviere et al., 2000).

3.5 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 represented as an aggregated alternative 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 the car with different departure time options. 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 options available to each user without any loss of significant 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. Railway system of Tokyo provides a thorough coverage in terms of both space and time. The frequency of rail in the morning is so high that all the rail users can choose a rail which allows them to reach their destination without any significant 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	

3.6 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 very basic structure tested is a MNL model assuming that no correlation exists between any of the alternatives. The nesting structure as well as covariance matrix form of this formulation is as shown in figure 1.

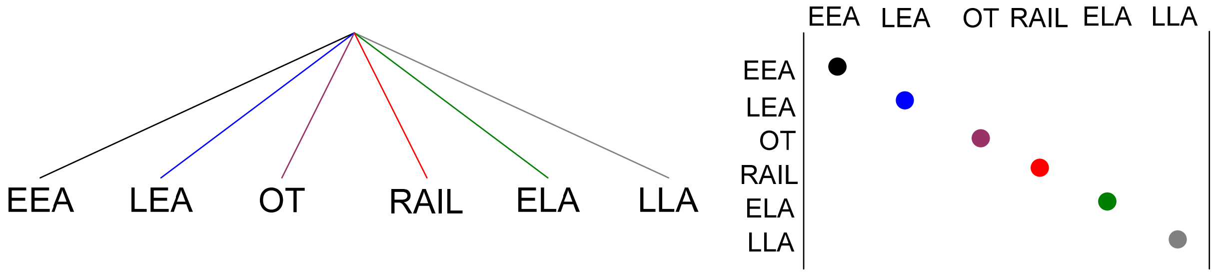

Figure 1 Multinomial Logit correlation structure and relevant covariance matrix

Figure 2 shows a nested logit model 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.

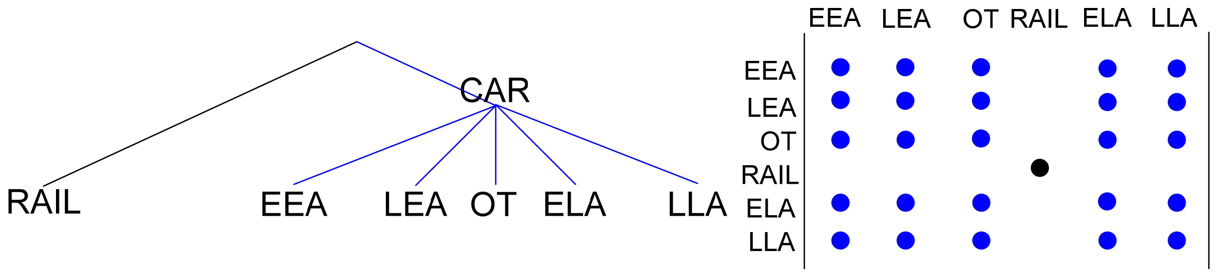

Figure 2 Nested Logit (mode based) correlation structure and relevant covariance matrix

Figure 3 shows a nesting structure in which all the alternatives are assigned to three nests. Rail is in a separate nest i.e. is not correlated with any other alternative while the departures using car resulting in early or on-time arrival at the destination are grouped in one nest and 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.

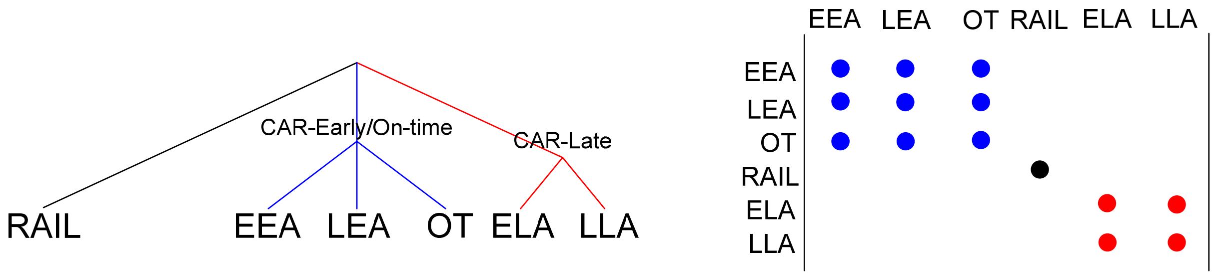

Figure 3 Nested Logit (3nests-a) correlation structure and relevant covariance matrix

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 corresponding to early, on-time and late arrivals. Each nest has two alternatives. The corresponding covariance matrix is also shown in the figure 4.

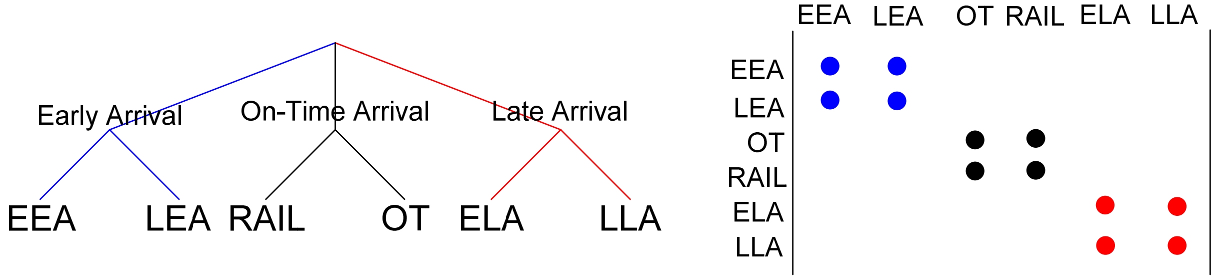

Figure 4 Nested Logit (3nests-b) correlation structure and relevant covariance matrix

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.

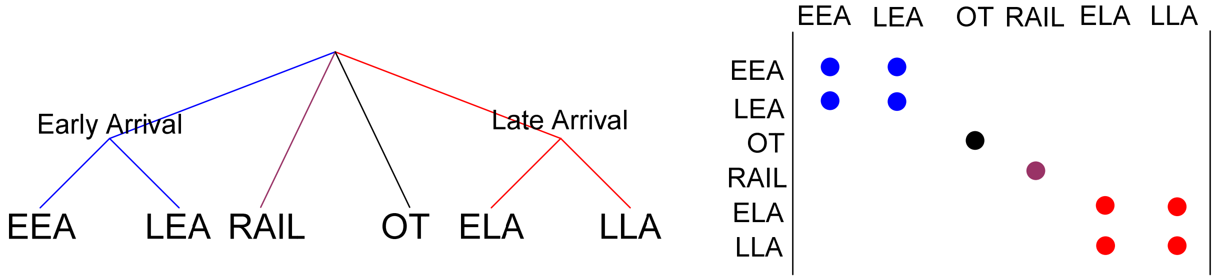

Figure 5 Nested Logit (4nests) correlation structure and relevant covariance matrix

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

ASC_{Rail} : Alternative Specific Constant for rail

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

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

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

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

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

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

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

$\beta_{\text{old age}}$: 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

$\beta_{\text{high income}}$: 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

$\beta_{\text{work in suburbs}}$: 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

$\sigma_{\text{work in suburbs}}$: Covariance 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

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

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

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

$\xi_{\text{early arrival/on time}}$: 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 the closed-form GEV and mixed GEV models respectively. These methods try to maximize the log-likelihood function. As stated earlier, the data 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

The first model structure estimated for combined mode and departure time is the MNL model as depicted in figure 1. Different utility specifications were tested to find the best possible utility function explaining the maximum variance in the data. Two alternate utility specifications are mentioned in table 1. One only including the level of service attributes whiles other also including the personal characteristics of the commuters. The model including the personal characteristics 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. Personal characteristics chosen for the inclusion in the model are those which have been found to be significant during different trials to affect the model performance. The level of service attributes only model is good for the network wide applications where the detailed personal characteristics data of the commuters is not available.

The utility functions for the LOS only model is,

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

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

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

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

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

$$V_{Rail} = ASC_{Rail} + \beta_{travel\,time}.TT_{Rail} + \beta_{\cos t}.COST_{Rail}$$

While for the level of service as well as socio-demographic attributes, the best possible utility function specification is,

$$V_{car,EEA} = \beta_{travel\,time}.TT_{EEA} + \beta_{\cos t}.COST_{EEA} + \beta_{early\,arrival}.EA_{EEA} + \beta_{late\,arrival}.LA_{EEA} + \beta_{car\,availability}.car\,availability + \beta_{high\,income}.high\,income$$

$$V_{car,LEA} = \beta_{travel\,time}.TT_{LEA} + \beta_{\cos t}.COST_{LEA} + \beta_{early\,arrival}.EA_{LEA} + \beta_{late\,arrival}.LA_{LEA} + \beta_{car\,availability}.car\,availability + \beta_{high\,income}.high\,income$$

$$V_{car,OT} = \beta_{travel\,time}.TT_{OT} + \beta_{\cos t}.COST_{OT} + \beta_{early\,arrival}.EA_{OT} + \beta_{late\,arrival}.LA_{OT} + \beta_{car\,availability}.car\,availability + \beta_{high\,income}.high\,income$$

$$V_{car,ELA} = \beta_{travel\,time}.TT_{ELA} + \beta_{\cos t}.COST_{ELA} + \beta_{early\,arrival}.EA_{ELA} + \beta_{late\,arrival}.LA_{ELA} + \beta_{car\,availability}.car\,availability + \beta_{high\,income}.high\,income$$

$$V_{car,LLA} = \beta_{travel\,time}.TT_{LLA} + \beta_{\cos t}.COST_{LLA} + \beta_{early\,arrival}.EA_{LLA} + \beta_{late\,arrival}.LA_{LLA} + \beta_{car\,availability}.car\,availability + \beta_{high\,income}.high\,income$$

$$V_{Rail} = ASC_{Rail} + \beta_{travel\,time}.TT_{Rail} + \beta_{\cos t}.COST_{Rail} + \beta_{old\,age}.old\,age + \beta_{young}.young + \beta_{work\,in\,suburbs}.work\,in\,suburbs$$

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 confidence 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.

Table 2 Estimation Results of MNL model of Fig. 1

Coefficients	Level of Service Attributes		Level of Service + Socio-demographic Attributes	
ASC_{Rail}	0.7040	(5.13)	1.8450	(9.38)
$\beta_{travel\ time}$	-0.0300	(-3.84)	-0.0314	(-3.91)
β_{cost}	-0.0008	(-3.15)	-0.0008	(-3.21)
$\beta_{early\ arrival}$	-0.0167	(-7.03)	-0.0180	(-7.28)
$\beta_{late\ arrival}$	-0.0844	(-7.78)	-0.0883	(-7.92)
$\beta_{car\ availability}$			0.9535	(6.64)
$\beta_{old\ age}$			1.3211	(2.48)
$\beta_{high\ income}$			1.1410	(4.20)
β_{young}			-0.5656	(-2.91)
$\beta_{work\ in\ suburbs}$			-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

A positive alternative specific constant for the rail mode 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 remaining traffic demand and a good spatial and temporal coverage provided by the railway network. Positive values for $\beta_{car\ availability}$ and $\beta_{high\ income}$ indicate that people owning a car or having higher incomes prefer to use car as their mode of choice as expected. A positive $\beta_{old\ age}$ in rail utility function indicates that old people prefer to use railway over car which is as expected. A negative value of $\beta_{work\ in\ suburbs}$ in the rail utility function indicates that people working in suburbs prefer to use car over the rail. This is quite understandable owing 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 NL Models

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

The results of these four nested logit models are reported in table 3. The results were estimated using the MNL parameters as initial values and different runs with different initial values indicate that a stable solution is obtained. All the parameters are significant in all the four models at more than 95% significance level except the $\beta_{old\ age}$ 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 are 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 savings is different from the MNL model for 2-nest model; it is higher while in other nested models it is lower than the MNL model. Values for early arrival and late arrival penalties are also resilient to the changes in correlation structure.

No clear winner among these alternate nesting structures can be established as all the nesting structures are significant although the nesting structure which divides the alternatives into three nests of early arrival, on-time arrival and late arrival perform marginally better than other models.

Table 3 Estimation Results of NL models of Fig. 2, 3, 4, 5

Coefficients	2 Nest NL Model (Fig. 2)		3 Nest NL Model (Fig. 3)		3 Nest NL Model (Fig. 4)		4 Nest NL Model (Fig. 5)	
ASC_{Rail}	2.7988	(5.04)	2.9832	(5.24)	5.1608	(3.66)	5.0572	(3.25)
$\beta_{travel\ time}$	-0.0414	(-3.78)	-0.0495	(-3.48)	-0.0785	(-3.54)	-0.0767	(-3.15)
β_{cost}	-0.0010	(-3.30)	-0.0014	(-3.22)	-0.0025	(-3.53)	-0.0020	(-3.07)
$\beta_{early\ arrival}$	-0.0231	(-6.51)	-0.0309	(-4.63)	-0.0463	(-4.68)	-0.0458	(-3.99)
$\beta_{late\ arrival}$	-0.1026	(-7.15)	-0.1587	(-4.31)	-0.2616	(-4.04)	-0.2455	(-3.60)
$\beta_{car\ availability}$	1.4848	(4.54)	1.5469	(4.68)	2.8604	(3.43)	2.7300	(3.08)
$\beta_{old\ age}$	2.0729	(2.36)	2.2378	(2.20)	3.9534	(2.00)	3.7227	(1.92)*
$\beta_{high\ income}$	1.7708	(3.41)	2.0178	(3.44)	3.4333	(2.94)	3.2418	(2.75)
β_{young}	-0.8750	(-2.65)	-1.0335	(-2.47)	-1.6697	(-2.29)	-1.6117	(-2.18)
$\beta_{work\ in\ suburbs}$	-1.6012	(-4.74)	-1.6217	(-5.54)	-3.1012	(-3.58)	-2.9553	(-3.28)
$\mu(0)$	0.6275	(5.67)	0.5486	(4.67)	0.3254	(3.98)	0.3461	(3.49)
(1)		(3.37)		(-3.84)		(-8.26)		(-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		-977.4		-969.8		-974.6	
Rho-Squared	0.327		0.328		0.333		0.330	
Rho-Squared bar	0.319		0.321		0.326		0.322	
VTTS(yen/min)	42.2		35.8		31.1		37.7	
VEAP(yen/min)	23.5		22.4		18.4		22.5	
VLAP(yen/min)	104.5		114.9		103.8		120.6	

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, different cross-nesting structures were tried and two of them which have been found to be highly significant are shown in figure 6.

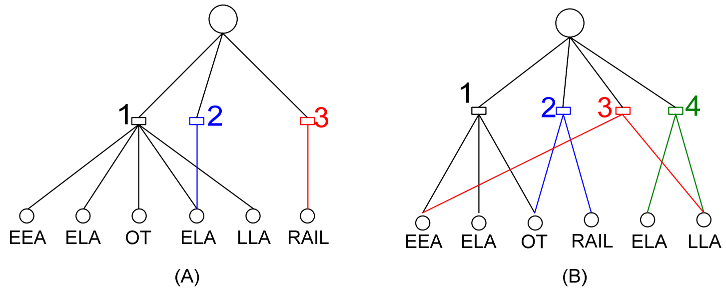

Figure 6 Two Cross-Nesting Structures

Cross-nesting structure (A) shows that Earliest Late Arrival Option belongs to two nests; one corresponding to car mode (nest 1) and another independent nest (nest 2). Rail is found to be independent of any correlation with other mode or departure time alternatives. Cross-nesting structure (B) 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 options 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. 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) respectively.

Values of travel time savings for 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 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 and confirms the fact that different correlation structures co-exist among different alternatives as indicated previously by different significant nesting structures in section4.2.

Table 4 Estimation Results of Cross-nested Logit Models

Coefficients	Cross-Nesting Structure (A)		Cross-Nesting Structure (B)	
ASC_{Rail}	8.2783	(3.7)	1.037	(3.15)
$\beta_{travel\ time}$	-0.0789	(-3.92)	-0.1692	(-2.8)
β_{cost}	-0.0028	(-4.26)	-0.0046	(-3.28)
$\beta_{early\ arrival}$	-0.0531	(-6.28)	-0.0830	(-3.82)
$\beta_{late\ arrival}$	-0.2738	(-5.46)	-0.516	(-3.57)
$\beta_{car\ availability}$	4.5459	(3.53)	5.5140	(3.10)
$\beta_{old\ age}$	6.4152	(2.22)	7.7750	(2.13)
$\beta_{high\ income}$	5.4582	(2.85)	6.7083	(2.68)
β_{young}	-2.6564	(-2.34)	-3.0507	(-2.12)
$\beta_{work\ in\ suburbs}$	-4.9020	(-3.71)	-5.9807	(-3.22)
μ (0)	0.2003	(4.11)	0.166	(3.58)
(1)		(-11.64)		(-17.95)
$\alpha_{ELA,1}$ (1)	0.606	(-4.04)	--	--
$\alpha_{ELA,2}$ (1)	0.394	(-6.21)	--	--
$\alpha_{EEA,1}$(1)	--		0.762	(-2.6)
$\alpha_{EEA,3}$(1)	--		0.238	(-8.3)
$\alpha_{OT,1}$(1)	--		0.745	(-3.8)
$\alpha_{OT,2}$(1)	--		0.255	(-11.1)
$\alpha_{LLA,3}$(1)	--		0.464	(-3.5)
$\alpha_{LLA,4}$(1)	--		0.536	(-3.04)
No. of observations	1324		1324	
No. of parameters	12		14	
Null-log likelihood	-1454.56		-1454.56	
Final-log likelihood	-959.3		-958.56	
Rho-Squared	0.340		0.341	
Rho-Squared bar	0.332		0.329	
VTTS(yen/min)	28.2		36.8	
VEAP(yen/min)	19.0		18.0	
VLAP(yen/min)	97.8		112.2	

4.4 Random Coefficient Multinomial Logit Models

As discussed before that 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. Several trial runs of the model using the random coefficient models yielded the results that late arrival as well as work in suburbs are the variables which have significant random coefficients. It is assumed that these random coefficients are distributed normally.

Results of random coefficient estimation for different random coefficient specifications are shown in Table 5. First mixed MNL (MMNL) 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 ($2(L(MMNL)\text{-}L(MNL))$ = 3.0 > 2.706, 90^{th} 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 built using work in suburbs as well as the late arrival as random variables. All the parameters are significant at the 95% significant 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% significant level. This indicates important gains in model performance obtained by using the random coefficient models. 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 of 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 in 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.

Table 5 Estimation Results of Mixed MNL model of Fig. 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)
$\beta_{travel\ time}$	-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)
$\beta_{early\ arrival}$	-0.0198	(-6.94)	-0.0273	(-8.30)	-0.0217	(-8.63)
$\sigma_{early\ arrival}$	--	--	--	--	0.0011	(2.04)
$\beta_{late\ arrival}$	-0.0936	(-7.65)	-0.3023	(-5.07)	-0.2545	(-5.10)
$\sigma_{late\ arrival}$	--	--	-0.17	(2.99)	-0.1504	(-4.65)
$\beta_{car\ availability}$	1.0421	(6.34)	1.1042	(5.86)	0.8612	(5.96)
$\beta_{old\ age}$	1.5960	(2.52)	2.0469	(2.99)	0.5516	(1.53)*
$\beta_{high\ income}$	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)*
$\beta_{work\ in\ suburbs}$	-0.9832	(-6.13)	-0.9942	(-5.16)	-1.013	(-7.50)
$\sigma_{work\ in\ suburbs}$	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		-1454.56	
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	

Third model described in table 5 shows the results for mixed logit estimation where early and late arrival as well as work in suburb parameters are normally distributed. Results indicate a decrease in the significance level for some of the parameters below 95% confidence 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.

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 4.4 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 by fitting a mixed nested logit model to the dataset. 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 6 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 confidence level for two of the nesting structures.

Comparison of the estimated model results with the corresponding MMNL model indicates an improvement in the fitness of the model 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 MMNL model in the previous section.

Similar to the NL models shown in section 4.2, no nesting structure is a clear winner over the others while all provide better fit over the corresponding MMNL and MNL models. Similar cross-nesting structures as shown in section 4.3 do not result in significant improvements in log-likelihood values though all the parameters are significant. These results are not reported here.

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.

Table 6 Estimation Results of Mixed NL models of Fig. 2, 3, 4, 5

Coefficients	2 Nest NL Model (Fig. 2)		3 Nest NL Model (Fig. 3)		3 Nest NL Model (Fig. 4)		4 Nest NL Model (Fig. 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)
μ(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 Mean (yen/min)	233.3		286		265		287	
VLAP Variance (yen/min)	144.4		169		157		169	

4.6 Error Component Models

Mixed logit models can also be utilized from another interpretation viewpoint and that is to model the correlation structures among the alternatives. It is usually said that error component logit models can virtually approximate any other modelling structure for discrete choices provided appropriate specifications are indicated. This section describes four error component logit models corresponding to four nested logit structures depicted in section 3.4 and used in section 4.2 and 4.5. Random error terms are introduced across alternatives which remain constant for a given set of alternatives hence, indicating the correlation among them.

Table 7 Estimation Results of Error Component Logit models

Coefficients	Corresponding to 2 Nest NL Model (Fig. 2)		Corresponding to 3 Nest NL Model (Fig. 3)		Corresponding to 3 Nest NL Model (Fig. 4)		Corresponding to 4 Nest NL Model (Fig. 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)
$\xi_{on\ time}$					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)*
$\xi_{early\ arrival/on\ time}$			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	

Table 7 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 high significance level. This shows that error component logit models can better capture the correlation structures as compared to the nested logit models.

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 $\beta_{\text{old age}}$ 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 some interaction among the late arrival parameters and the error components and needs further investigation.

Comparison of these error component logit models with the mixed logit model shows them at par with each other but inferior to the nested mixed 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.

4.7 Overview of the estimation results

The overall model estimation results are summarized in Table 8 as follows. The table presents the final log-likelihood values and the number of estimated parameters for all the runs performed above.

First, inclusion of both socio-demographic and level of service variables greatly improves the model fit compared to a model with only level of service variables. 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. Based on the fact that all the four nesting structures were found to be significant, different cross-nesting structures were tried and two structures were found to show significant improvement in the model performance. 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 as well as work in suburbs were found to be having significant random coefficients.

Taste heterogeneity was found to be significant in case of work in suburbs attribute where the distribution of the parameter was such that about 33% of the commuters will have a positive utility for using trains to work while in the case of the MNL model everyone was having a negative utility when travelling to work in suburbs.

Table 8 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 Fig. 2)	-978.9	11
NL_2 (structure defined in Fig. 3)	-977.4	11
NL_3 (structure defined in Fig. 4)	-969.8	11
NL_4 (structure defined in Fig. 5)	-974.6	11
Cross Nesting Structure (A)	-959.3	12
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

Taste heterogeneity also indicated distributed value of late arrival penalty across the individuals, indicating differences between the commuters. This fact may 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 structure of nested logit models and mostly perform better than corresponding nested logit models. However, these models did not outperform the Mixed Nested Logit models.

5. Conclusions

A stated choice survey of departure time choice of morning commuters under schedule constraints was conducted in the Tokyo Metropolitan Area. Data collected in this survey was used for the estimation of a combined mode and departure time choice model for the morning commuters in Tokyo area.

Different model specifications accounting for the correlation among different alternatives as well as random taste variations across individuals were tried in this study. Results indicate that accounting for such phenomenon provide a better fit model than the ordinary MNL models which assume no correlation among the alternatives and suffer from the independence of irrelevant alternative (IIA) property especially in the context of departure time choice models where the consecutive alternatives are highly correlated.

Nested Logit models perform better than the MNL by accounting for the correlation among alternatives while the MMNL performs better than the MNL by accounting for the random taste variations. Mixed Nested Logit models perform better than NL models as well as MMNL models by accounting for both correlation structure as well as random taste variations jointly. Error component logit corresponding to the correlation structures of the nested logit models were developed and have been found to mostly perform better than corresponding NL models.

It is noticeable that with respect to the four proposed nesting structures, no correlation structure was found to perform clearly better than other structures. This indicates the possibility of different alternatives belonging to more than one nest. Further analysis using the cross-nested logit 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 is found that CNL model can better accommodate the similarity among alternatives than the NL model, while keeping a closed-form probability function. This may be important for future applications of the proposed models because Mixed Logit models require simulations which can be computationally expensive while simple MNL, NL or CNL can be easily computed. Results of CNL models also indicate that departure time intervals are not only correlated by their proximity to each other but the departure time intervals far from each other but having similar schedule delays are also correlated.

Results show that the late arrival penalties for the commuters are much 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. The first application of this model will be to evaluate the multimodal transportation network in the Tokyo Metropolitan Area, and the results will be published in subsequent papers.

References

Abkowitz, M. (1981). An Analysis of the Commuter Departure Time Decision. *Transportation* **10**, 283-297.

Ben-Akiva, M. and M. Bierlaire (1999). Discrete choice methods and their applications to short-term travel decisions, In: *Handbook of Transportation Science* (edited by R. Hall), International Series in Operations Research and Management Science, Vol. 23.

Bhat, C. (1998a). Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research 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 B* **32** (7), 455–466.

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

Bierlaire, M. (2005). An introduction to BIOGEME (Version 1.4).

http://biogeme.epfl.ch/doc/tutorial.pdf.

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

Cardell, N. and F. Dunbar (1980). Measuring the societal impact of automobile downsizing. *Transportation Research A* **14** (5–6), 432–434.

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

Chin, A. (1990). Influences on Commuter Trip Departure Time Decisions in Singapore. *Transportation Research A*, **24** (5), 321-333.

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

de Palma, A., M. Ben-Akiva, C. Lefevre, and N. Litnas (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 G. Kocur (1981). Schedule Delay and Departure Time Decisions in a Deterministic Model. *Transportation Science*, **15**, 62-77.

Hendrickson, C., D. Nagin, and E. Plank (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.

Hendrickson, C. and E. Plank (1984). The Flexibility of Departure Times for Work Trips. *Transportation Research A*, **18** (1), 25-36.

Hess, S., M. Bierlaire, and J. Polak (2004). Development and Application of a Mixed Cross-Nested Logit model. European Transport Conference, Strasbourg.

Hurdle, V. (1974). The effect of queuing on traffic assignment in a simple road network. Proceedings of the 6th International Symposium on Transportation and Traffic Theory. Elsevier Science, Amsterdam.

Hurdle, V. (1981). Equilibrium Flows on Urban Freeways. *Transportation Science* **5**, 255-293.

Koppelman, F. S. and C.-H. Wen (2000). The Paired Combinatorial Logit model: properties, estimation and application. *Transportation Research* **34**B, 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 G. Newell (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, 21-40.

Louviere, J., D. Hensher, and J. Swait (2001). *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: *Spatial Interaction Theory and Residential Location* (Edited by A. Karlqvist et al.), North-Holland, Amsterdam.

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

Papola, A. (2004). Some developments on the cross-nested Logit model. *Transportation Research* **38** B(9), 833–854.

Polak, J. W. and P. M. Jones (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** (3), 467-479.

Small, K. (1987). A discrete choice model for ordered alternatives. *Econometrica* **55**(2), 409-424.

Shimizu, T. and T. Yai (1999). An Analysis of The Effect On Peak Period Toll in Tokyo Metropolitan Expressway. *Journal of the Eastern Asia Society for Transportation Studies*, **3** (5), 305-318.

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

Vickrey, W. (1969). Congestion theory and transport investment. *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 S. Bekhor (1998). The Link-Nested Logit model of route choice: overcoming the route overlapping problem. *Transportation Research Record* **1645**, 133–142.

Walker, J. L. (2001). *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables*. PhD thesis, MIT, Cambridge, MA.

Wang, J. (1996). Timing utility of daily activities and its impact on travel. *Transportation Research A* **30** (3), 189–206.

Wen, C.-H. and F. S. Koppelman (2001). The Generalized Nested Logit Model. *Transportation Research* **35** B (7), 627–641.