Development and Calibration of Route Choice Utility Models: Factorial Experimental Design Approach

Yaser E. Hawas[1]

Abstract: This paper provides a detailed overview on the state of practice in the area of data collection, modeling, calibration, and validation of route choice models. This paper presents a new approach for modeling route choice behavior using stated preference data coupled with rating (scoring) technique. The data collection was carried out over three stages; travelers' scores of route attributes, attributes' levels identification, and experimental templates for validation. A factorial experiment-design model was developed and validated through a two-stage procedure: the validation of the scoring technique, and the validation of the developed factorial experiment design model. Results indicated the dominant influence of travel time, and queuing time information on the route choice decisions. Other factors such as the speed, highway class, familiarity, and highway/pavement conditions were also found significantly important. The results include the percentages of the factors' contributions to the route utility, and the relationships derived to illustrate the effect of the factors' interactions on the travelers' route utility.

DOI: 10.1061/(ASCE)0733-947X(2004)130:2(159)

CE Database subject headings: Route preference; State-of-the-art reviews; Questionnaires; Data collection; Statistical models; Factor analysis; Calibration; Validation.

Introduction

Route choice behavior models represent the core of the simulation models to support the advanced traveler information systems (ATIS). With the many developments of on-line descriptive information systems for route guidance, there is a noticeable deficiency in modeling the drivers' response to these information systems. There are several difficulties in dealing with route choice behavior modeling. First, the variance in the travelers' perceptions of the provided information; a level considered acceptable by one traveler, might be regarded unpleasant by another. Second, the variation of the travelers' appraisal to the various route attributes. While the majority of travelers may report the travel time as the common factor affecting their route selection, some travelers may appraise other factors such as the familiarity of the route, and its pavement condition. Third, the travelers' compliance to the advisory messages may also differ over time as a result of the reliability of information and actual driver's encounters. Finally, the travelers may use various criteria to identify the *best* route; identifying these criteria (although challenging) is essential for the calibration of the route choice models.

Laboratory experiments (either through simulation experiments, driving simulators, internet graphical simulators, interactive surveys) were recognized as effective tools for data collection (Mahmassani and Herman 1990; Bonsall and Parry 1991; Adler 1993; Vaughn et al. 1993; Koutsopoulos et al. 1994; Lotan 1997; and Ozbay et al. 2001). Koutsopoulos et al. (1995) examined various traveler simulators as a source of data for driver's behavior. In light of the deficiencies (biasness) that may be incurred due to the laboratory nature, recommendations for simulators design are discussed. Conventional data surveying methods such as phone interviews or mail surveys were also extensively used (Abdel-Aty et al. 1994; Polydoropoulu et al. 1994; Abdel-Aty et al. 1995a; Katsikopoulos et al. 2000).

Day-to-day route choice behavior was studied by Chen and Mahmassani (1993); Srinvasan and Mahmassani (2000); and Ozbay et al. (2001). Chen and Mahmassani (1993) developed an interactive multiuser simulator to capture the day-to-day switching behavior in response to travel-time information. Drivers' decisions are explicitly accounted for in updating the network congestion status. Srinvasan and Mahmassani (2000) formulated the route choice compliance as a sequence of compliance and inertia (tendency of users to adhere to their usual current paths). A nested multinomial probit model was developed. Ozbay et al. (2001) developed a day-to-day route choice model using data from an Internet-based graphical simulator.

The effect of information accuracy, route familiarity, and compliance was studied extensively by Bonsall and Parry (1991); Vaughn et al. (1993); Abdel-Aty et al. (1994); and Lotan (1997). Bonsall and Parry (1991) developed an interactive simulation model that displays at each junction the *best* route with variation in information accuracy. The degree of compliance was measured (Bonsall 1992). Bonsall et al. (1997) discussed the methodology to calibrate the route choice simulator on familiar and nonfamiliar drivers. Response to variable message signs (VMS) information was also tested and validated.

Vaughn et al. (1993) proposed a computer-based simulation to collect route choice data. The experiments were designed to capture the pretrip route choice behavior under various levels of advisory information accuracy. Analysis of variance was performed and a binary logit model was developed. Results indicated a

[1]Associate Professor, Civil and Environmental Engineering Dept., UAE Univ., P.O. Box 17555, Al-Ain, United Arab Emirates. E-mail: y.hawas@uaeu.ac.ae

Note. Discussion open until August 1, 2004. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on May 16, 2002; approved on March 6, 2003. This paper is part of the ***Journal of Transportation Engineering***, Vol. 130, No. 2, March 1, 2004. ©ASCE, ISSN 0733-947X/2004/2-159–170/$18.00.

threshold of accuracy, below which drivers do not comply. The prediction (of drivers' response) accuracy of the logit model is estimated to be about 80%. Abdel-Aty et al. (1994) utilized a computer-aided telephone interview for data collection. Cross tabulation was performed to explore the interrelation among the variables affecting pretrip route choice. Two sets of models were estimated: bivariate probit and negative binomial. The models are used to predict the frequency of route change per month. Vaughn et al. (1995a) conducted simulation experiments under various driving scenarios to test the effect of recurrent and nonrecurrent information. Analysis of variance and regression techniques indicated the significance of incident-related information. Vaughn et al. (1995b) developed a dynamic linear probabilistic compliance frequency model.

Abdel-Aty et al. (1995a) presented a statistical analysis of commuters' route choice extracted from mail survey. A binary logit model was developed using stated preference repeated-measurement data. The results indicated the significance of travel time, information reliability, safety, and roadway characteristics on the route choice. Abdel-Aty et al. (1995b) reported the various surveying methods for route choice modeling. Abdel-Aty et al. (1997) developed two binary logit models using computer-aided phone interview (Abdel-Aty et al. 1994) and a mail survey route choice data (Abdel-Aty et al. 1995a).

Lotan (1997) studied route choice behavior for familiar and nonfamiliar drivers. Data collection was carried out via a driving simulator. Unfamiliar drivers illustrated more homogeneity in their switching behavior than familiar drivers did. Two models were developed; an approximate reasoning-based and a random utility model. The key contribution of this work was the development of a procedure to model how unfamiliar drivers become familiar. Lotan (1998) presented a two-stage framework for modeling discrete choice behavior.

The effect of the information media was studied by Polydoropoulu et al. (1994); Emmerink et al. (1996); Peeta et al. (2000); and Katsikopoulos et al. (2000). Polydoropoulu et al. (1994) developed a model (using school trip data) for the acquisition media, pretrip and en-route information, drivers' route switching behavior, and willingness to pay for accurate information. They concluded the importance of the attitudinal factors and acquisition media on en-route diversion. Emmerink et al. (1996) analyzed the impact of both radio and variable message sign information on route choice behavior. Several discrete choice models were developed: ordered probit, multiple logit, and bivariate ordered probit. Among the findings is that route choice is strongly related to the type and distance of the alternative roads. They concluded similar effect of both radio and VMS. Peeta et al. (2000) developed a set of logit models for VMS information, utilizing data collected from on-site stated preference user survey. They concluded that the level of information details, socioeconomic characteristics, knowledge of the network (familiarity) are among the important factors influencing the VMS compliance rates. Katsikopoulos et al. (2000) conducted two experiments on same participants; first, using simplified paper drawings, and second using a media of virtual reality. The willingness of switching was measured, and a simple probabilistic model was developed to describe the route switching behavior for the two experiments.

The majority of the above approaches are discrete choice models. These models are limited by the inability to capture the subjects' taste variation, and the vagueness of the subject behavior. The independence of irrelevant alternatives property of the multinomial logit model imposes the restriction of zero covariance between the utilities of alternatives, and this restriction is inappropriate in many choice situations (Koppleman and Wen 2000). The use of the nested logit model would relax the zero covariance restriction but will impose the restriction of equal covariance among all alternatives in a common nest and zero covariance elsewhere. Koppleman and Wen (2000) recently introduced the paired combinatorial logit model for relaxation of the restrictions. The premise of these models is the ability to account for taste variations (Koppleman and Wen 2000; Sermons and Koppleman 1998; Swait and Bernardino 2000). In brief, it could be claimed that the recent discrete choice models are capable of incorporating the subjects' taste variation.

The problem then remains with modeling the vagueness (fuzziness) in the subject behavior. Kikuchi and Pursula (1998) discussed the treatment of uncertainty in transportation systems. They classified the uncertainty in the drivers' behavior and perception into two classes: vagueness (in distinguishing the transition between states) and ambiguity (when no evidence or data exist to support the truth of a proposition). The fuzzy set theory was suggested for the applications that entail vagueness. Recently, *pure* fuzzy logic was recognized as a modeling tool for route choice (Koutsopoulos et al. 1994; Teodorovic et al. 1998; Henn 2000; Kuri and Pursula 2001; Rilett and Park 2001). It alleviates the limitations of the distribution functions of the conventional discrete choice models. However, developing the knowledge base of the fuzzy logic is conventionally done intuitively using simple reasoning rules. As such, very limited size knowledge bases with few factors (primarily travel time) were considered. Several factors of potential influence on route choice were ignored to simplify the knowledge bases. Furthermore, no specific methodology is available for the systematic calibration of such knowledge bases. The memberships (fuzzy sets' parameters) and the knowledge base (rules) are conventionally set intuitively using reasoning arguments of huge data sets or trail and error. In complex systems, this intuitive reasoning is at best a challenging task, if not impossible.

The most notable pure fuzzy-logic-based route-choice models are those by Koutsopoulos et al. (1994); Teodorovic et al. (1998); Henn (2000); and Kuri and Pursula (2001). Koutsopoulos et al. (1994) utilized a PC-based driving simulator to develop route choice models using fuzzy sets and approximate reasoning. Teodorovic et al. (1998) presented a fuzzy logic model for route choice. The knowledge base was developed using an approximate reasoning algorithm, with computer simulation data. Henn (2000) accounted for uncertainties of the drivers' behavior. The reliability effect of the ATIS information was modeled as a modification of the probability that the traveler perceives regarding the predicted route cost. Kuri and Pursula (2001) compared between logit-type random utility models and fuzzy logic. Nondynamic stated preference choices were utilized to assess the effectiveness of the two models. The factors considered included travel time, cost, roadway type, and delays. Their work proved the fuzzy logic to be more effective in capturing the taste variation. Rilett and Park (2001) focused on the incorporation of uncertainty and multiple objectives in real-time route selection using fuzzy logic.

Fuzzy logic can be used effectively in addressing the two issues of taste variation and vagueness, i.e., capturing the variability of the travelers' appraisal to the different route attributes, as well as the variability in their perceptions to the various attribute levels. To overcome the limitations encountered in developing *pure* fuzzy logic systems (intuitive reasoning to develop the knowledge base), the neuro-fuzzy logic (integrated fuzzy and neural nets) may be used. The idea is to couple the fuzzy logic with ability to replicate complex systems' behavior by utilizing the training capabilities of the neural nets. The training is used to calibrate the fuzzy sets, parameters, and knowledge base with the objective of replicating the logic of route choice actually undertaken by travelers.

The literature on the data collection for route choice modeling can be divided into three approaches: stated preference rating (scoring), stated preference ranking, and stated choice. In the first approach, the subjects are asked to rate the various routes according to the perceived route utility. The second approach requires only ranking, and the third depends on the observation of the subjects' actual choices. The first approach may be criticized for the vagueness in the subjects' definitions of the route utility. The route utility rating data may not be accurate if the subject does not have a clear definition of the route utility. Nonetheless, the rating methodology could be quite successful if the rating is done for the tangible route attributes (such as travel time, speed, and pavement condition). The stated preference choice data may be more accurate, but they require enormous resources-usually carried out via real-life experiments or interactive simulators. This paper adopts the approach of stated preference rating for the route attributes. Future work will focus on developing the route choice models using the other data types.

This paper focuses on developing a factorial experiment design model (FEDM) for the route utility estimation. This modeling approach may be utilized in assessing the significance of the various route attributes. It may also be utilized in studying the correlation and interaction effects among the various route attributes (rarely elaborated in the literature). The FEDM model is utilized afterwards to generate estimates (replicates) of the route utilities perceived by the travelers. These replicates are then used in Hawas (2004) to develop a neuro-fuzzy logic. The FEDM is based on the analogy that the traveler assigns specific rates or ''scores'' to the various route attributes according to the level of attribute he/she perceives, and then utilizes such scores to estimate an overall route utility.

This paper will only focus on the data collection and calibration of the FEDM for route utility. It is imperative to realize that the FEDM could be used to capture the inter-group taste variation, but neither the intra-group taste variation nor the vagueness in the choice behavior. It is based on the assumption that the travelers within the same socio-demographic group will exhibit the same behavior if subjected to the same attribute values. The fuzzy logic model (Hawas 2004) relaxes these restrictions. The fuzzy logic is calibrated using neural nets and the data produced by the FEDM developed herein.

This paper is divided into ten sections. The second section discusses the details of the attribute scoring technique and the travelers' data collection. The third section provides a preliminary statistical analysis of the travelers' survey data. The theoretical basics of the FEDM are reviewed in the fourth section. The scoring data preparation for the FEDM calibration is discussed in details in the fifth section. The sixth section summarizes the statistical results of the calibrated FEDM. The seventh section provides a description of the route factors' (and their interactions) parameters. The validation of the scoring technique and the FEDM is discussed in the eighth section. The ninth section provides detailed analysis of the FEDM results. The conclusions and future work are discussed in the last section.

Route Selection Behavior Modeling and Data Survey

As previously indicated, the discrete choice models are limited by inability to model taste variation (Koppleman and Wen 2000) and regions of stability and vagueness (Kikuchi and Pursula 1998) in the traveler's perception of route utility. The route utility associated with attribute i, u_i, is practically constant over the range $[x_i-\delta x_i, x_i+\delta x_i]$, where x_i is the value of attribute i, and δx_i is the value defining the range of stable route utility around x_i. To account for this issue, it is assumed that the route selection decisions are carried over by subjective analysis and evaluation. It starts by converting the numeric measures of the route attributes (traffic states) to ''linguistic'' terms—fuzzification process (Hawas 2004). For example, a travel time of 30 min is transferred to as ''medium'' or ''high'' travel time. The route decisions are then done subjectively using these linguistic descriptions, and finally converted again to numeric values by the so-called defuzzification process (Hawas 2004).

In this paper, the modeling approach partially accounts for the traveler's socio-demographic characteristics. Due to the constraints imposed by the resources for data collection and analysis, this paper considers only six traveler groups based on the common traffic characteristics in the residence area and age (will be discussed more in the following section).

The decision-making involves the utilization of a scoring technique. The evaluation of the routes accounts for the scores (relative weights) assigned by the traveler for the routes' attributes (factors) as well as their levels. Each attribute (factor) is categorized into various linguistic levels. For example, the pavement condition may be categorized as *poor*, *acceptable*, or *good* condition. Six factors of potential influence on the driver's decisions to undertake a specific route are considered herein. These factors were extracted via focus group discussions:

1. Travel time (along the route);
2. Highway classification;
3. Queuing time;
4. Familiarity (with the route);
5. Highway/pavement condition; and
6. Speed.

This paper is based on the idea of utilizing subjective analysis through the scoring (rating) of the route attributes (factors). Numerous scoring techniques were reported in the literature; Papacostas and Prevedouros (1993) provide detailed discussion on the various scoring techniques and applications. The details of the adopted scoring technique are discussed below.

Route Utility Modeling Approach: Scoring Technique

The modeling approach adopted in this paper assumes that the route utility perceived by the traveler is equal to the sum of the products of the attribute (factor) scores and the perceived level scores. This assumption is validated later in the eighth section. Given this assumption, the route utility can be expressed as follows:

$$U_i^{k,t}=\sum_{j=1}^{J}\sum_{\ell=1}^{L} S_j^k \hat{S}_{\ell,j}^k \delta_{i,j,\ell}^{k,t}, \quad \forall i\in I,\ k\in K_n,\ n=1,\ldots,N \tag{1}$$

where t=integer time index; i=route index, $i=1,\ldots,I$; j=attribute (factor) index, $j=1,\ldots,J$; ℓ=index of attribute's level perceived by the traveler, $\ell=1,\ldots,L$; k=participant (traveler) index; n=index of socio-demographic set (group) to which the participant belongs; K_n=set of participants in the nth group; $U_i^{k,t}$=absolute route utility assigned to route i by participant k at time t; S_j^k=score of factor j assigned by participant k; $\hat{S}_{\ell,j}^k$=score

Route Selection (Score Stage) Questionnaire

This questionnaire is intended for research in the area of traffic engineering. Your time and effort in addressing the questions are highly appreciated.

Socio-Demographic Section

Age: City:

Sex: Occupation:

Travel Preference Section

To characterize the drivers' route selection behavior, six variables of potential effect on route selection are identified as follows:

- Travel time of the highway
- Highway Class
- Speed
- Familiarity with the Highway
- Queue Time
- Highway/ Pavement Condition

On the attached table, write the appropriate "variable weights" to express to what extent your route selection is influenced by the above variables. Also, indicate using appropriate "term scores" your preference to undertake a highway exhibiting the term conditions (e.g. Freeway, High Travel Time, Medium Queue Time, Non-Familiar, etc).

Factor/Weight	Term	Term's Score
Travel Time /10	▪ Free	/10
	▪ Light	/10
	▪ Medium	/10
	▪ High	/10
Highway Classification /10	▪ Freeway	/10
	▪ Arterial (With Progression)	/10
	▪ Arterial (Without Progression)	/10
Queue Time /10	▪ Light	/10
	▪ Medium	/10
	▪ Heavy	/10
Familiarity /10	▪ Familiar	/10
	▪ Not Familiar	/10
Speed /10	▪ Low	/10
	▪ Medium	/10
	▪ High	/10
Highway/ Pavement Condition /10	▪ Poor	/10
	▪ Acceptable	/10
	▪ Good	/10

Fig. 1. Survey questionnaire for score stage

of level ℓ of factor j assigned by participant k; $\delta_{i,j,\ell}^{k,t}$ binary index with value of 1 if level ℓ of factor j is perceived by participant k along route i at time t, and 0 otherwise.

Note that the participants, even those within the same socio-demographic group, may not perceive the same linguistic level for a specific numeric traffic state. To calculate the relative route utility, $\tilde{U}_i^{k,t}$, of route i at time t, the following formula is used:

$$\tilde{U}_i^{k,t}=\frac{\sum_{j=1}^{J}\sum_{\ell=1}^{L}S_j^k\hat{S}_{\ell,j}^k\delta_{i,j,\ell}^{k,t}}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{\ell=1}^{L}S_j^k\hat{S}_{\ell,j}^k\delta_{i,j,\ell}^{k,t}}, \quad \forall i\in I,\ \ k\in K_n,\quad n=1,\ldots,N \tag{2}$$

The sum of the relative utility for all the routes is equal to unity

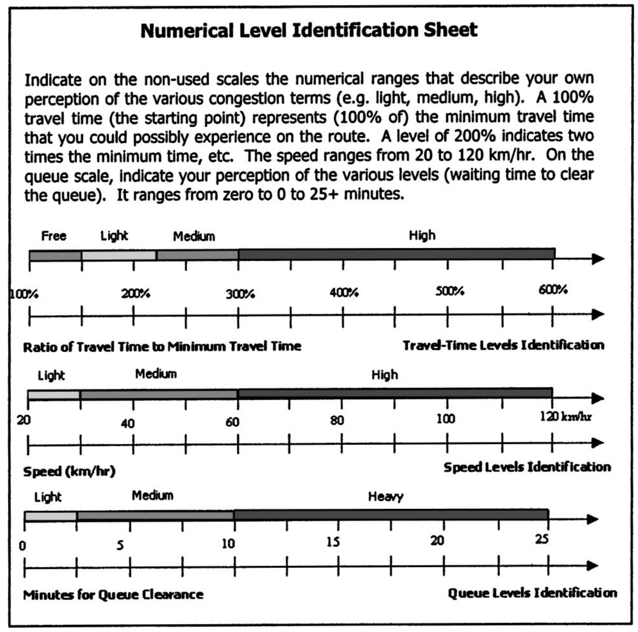


Fig. 2. Numerical levels' identification scale sheet

$$\sum_{i=1}^{I}\tilde{U}_i^{k,t}=1 \tag{3}$$

A traveler k will select the route exhibiting maximum absolute (or relative) utility. At time t, traveler k will select route m, where

$$\tilde{U}_m^{k,t}=\text{Max}\{\tilde{U}_i^{k,t}, i=1,\ldots,I\} \tag{4}$$

For the sake of simplicity in presenting, it is assumed that the factor and level scores are time independent. Nonetheless, the same approach can be adopted with dynamic scores.

Route Choice Behavior Surveying

The data needed for the FEDM calibration were collected using focus group discussions. It is imperative to stress that it is not the author's intention to draw any conclusive forms for the FEDM, but rather to illustrate the calibration and validation procedures. The FEDM (although calibrated from limited data) has proven to be quite effective in capturing the route choice behavior as will be explained later. The participants' groups were selected to partially capture the effect of socio-demographic characteristics on the choice behavior. The participants are grouped based on the age and the common traffic characteristics (congested, not congested) in their residence area. Travelers residing in congested areas may demonstrate more tolerance at high congestion levels than those residing noncongested areas, and as such, their scores (rating) of the congestion levels would differ. Similarly, old-age groups may demonstrate higher tolerance levels than young-age groups.

Various participants were targeted in the Al-Ain and Abu-Dhabi cities in UAE. While Al-Ain city commonly exhibits free-to-medium congestion conditions, Abu-Dhabi city suffers from severe congestion problems. Due to the time and resources' restrictions, the survey targeted only students, their parents, faculty members, and staff. Sixty participants were selected to cover the six city-age groups equally. Followed below is the list of targeted groups:

- NC-Y: noncongested city area and young age (<25 years);
- NC-M: noncongested city area and middle age (25–45 years);

Table 1. Preliminary Statistical Analysis of Score Stage Data

Factor	Overall factor score	NC-Y Mean	NC-Y S.D.	NC-M Mean	NC-M S.D.	NC-O Mean	NC-O S.D.	C-Y Mean	C-Y S.D.	C-M Mean	C-M S.D.	C-O Mean	C-O S.D.
Travel time	8.33	8.5	1.03	8.9	0.90	7.5	1.13	9.1	0.80	8.1	0.63	7.9	0.45
Class	5.85	5.9	0.96	7.3	0.45	6.2	0.75	4.3	1.20	5.4	0.90	6	0.96
Queue	8.26	6	0.63	8.2	0.50	8.6	0.70	8.4	0.45	9.4	0.70	9	0.45
Familiarity	5.95	5.3	0.90	6.8	1.05	7.9	0.95	3.4	0.60	5.3	0.50	7	0.80
Highway condition	4.92	4.1	0.56	5.3	0.63	6	0.90	3.2	1.50	4.2	1.63	6.7	0.83
Speed	7.33	7.4	0.93	6.9	0.83	5.2	1.05	8.9	0.93	8.7	0.80	6.9	0.90

(Average Factor Scores and Standard Deviations for Various Participant Groups)

- NC-O: noncongested city area and old age (45+ years);
- C-Y: congested city area and young age (<25 years);
- C-M: congested city area and middle age (25–45 years); and
- C-O: congested city area and old age (45+ years).

Group discussions were conducted with all participants; indicating the objectives, interpretation of factors and levels, etc. The survey was conducted over three stages; the score stage, the numeric levels identification, and the validation stage. In the score stage, participants were asked to assign scores to (rate) the factors and their levels, using the form shown in Fig. 1. The factor score (out of 10) represents the degree to which the participant's decisions are affected by such a factor. The level score reflects the preference to undertake a route exhibiting this particular level. The total route utility rating is equal to the sum of the products of the factor scores and the level scores [Eq. (1)].

In the numeric level identification stage, each participant was asked to indicate on the numeric scales shown in Fig. 2 his interpretation of the linguistic terms describing the factors' levels. The data collected in this stage are essential for defining the fuzzy sets needed for the calibration of the fuzzy logic (Hawas 2004). Each participant was asked to consider a route of a specific free-flow travel time (say 10 min), corresponding to the benchmark value of 100% on the travel time scale. The participant then indicates the intervals (in minutes) corresponding to his perception of the travel-time terms (e.g., free: [10,12], light: [12,15], etc.) Finally, the intervals are converted to the percentage scale shown in Fig. 2. Similarly, the queuing time and the speed terms are indicated on their scales. The purpose of the validation stage is to ensure the validity of the scoring technique and the effectiveness of the FEDM in replicating the route choice decisions.

Score Data: Preliminary Statistical Analysis

Table 1 shows the preliminary analysis of the factors' scores data. It indicates the average factor scores and standard deviations for the various groups. The NC-Y group is mostly sensitive to travel time and speed, NC-M is mostly sensitive to travel and queuing time, and the NC-O is mostly sensitive to queuing time and familiarity. The C-Y group, similar to NC-Y, is mostly sensitive to travel time and speed but with higher scores. The C-M group is mostly sensitive to queuing time and speed. The queuing time is the dominant factor for the C-O group. It can be easily concluded the travel time and queuing time are the factors of highest influence on the route choice decisions. While the highway class is important for the groups residing in noncongested areas, the familiarity and highway condition are important for old-age groups (NC-O and C-O). It can also be concluded that young-age groups are affected by the speed attribute more than the old-age groups. The group discussions were successful in highlighting some interesting behavioral choice attitudes; for example, some participants inclined preference towards *light* or *medium* (but not *free*) conditions; *free* conditions imply boring driving experience, particularly on long trips.

Route Utility Factorial Experimental Design Approach

The premise of the FEDM is its ability to quantify the effect of the factors, levels, and interactions on the route utility. The FEDM can be utilized to estimate the utility perceived by traveler k, at time t, as follows:

Table 2. Sample of Data for Factorial Experiment Design Model Calibration

City-age group (A)	Travel time (B)	Highway class[a] (C)	Queue time (D)	Familiarity (E)	Highway/pavement condition (F)	Speed limit (G)	Avg. group utility $\underline{U}^t_{i,g}$
NC-Y	Light	AWP	Heavy	Nonfamiliar	Good	Medium	**213.3**
NC-O	Free	Freeway	Medium	Familiar	Good	High	355.0
C-M	Medium	AWOP	Medium	Nonfamiliar	Acceptable	Low	217.7
NC-O	High	Freeway	Medium	Nonfamiliar	Good	Medium	232.2
C-O	High	Freeway	Light	Familiar	Good	High	291.3
NC-O	Medium	AWP	Heavy	Familiar	Good	Medium	226.2
C-Y	Free	Freeway	Medium	Familiar	Good	High	367.5
NC-O	High	Freeway	Heavy	Nonfamiliar	Acceptable	Low	138.9
NC-Y	Medium	AWOP	Medium	Nonfamiliar	Acceptable	High	249.4
C-O	Free	Freeway	Heavy	Familiar	Poor	Low	230.4
C-O	Medium	AWOP	Heavy	Familiar	Good	High	228.7

[a]AWP=arterial with signal progression and AWOP=arterial without signal progression.

Table 3. Results of Analyses of Variance on Factors and Interactions Effect

Factor or interaction	DOF	Sum of square of errors	Mean of square of errors	f Statistic	% Contribution
Intercept					
A	5	427,575	85,515	**335.080**	3.50
B	3	3,745,270	1,248,423	**4891.80**	30.69
C	2	963,169	481,584	**1887.03**	7.89
D	2	3,835,065	1,917,532	**7513.63**	31.42
E	1	768,831	768,831	**3012.57**	6.30
F	2	383,649	191,824	**751.642**	3.14
G	2	893,722	446,861	**1750.97**	7.32
AB	15	25,452	1,696	**6.648**	0.20
AC	10	18,800	1,880	**7.366**	0.15
AD	10	77,401	7,740	**30.329**	0.63
AE	5	10,111	2,022	**7.924**	0.08
AF	10	30,135	3,013	**11.808**	0.24
AG	10	26,687	2,668	**10.457**	0.21
BC	6	11,457	1,909	**7.482**	0.09
BD	6	780	130	0.509	0.006
BE	3	4,934	1,644	**6.444**	0.040
BF	6	297	49	0.194	0.002
BG	6	1,459	243	0.952	0.011
CD	4	6,294	1,573	**6.165**	0.051
CE	2	2,865	1,432	**5.613**	0.023
CF	4	473	118	0.463	0.003
CG	4	1,317	329	1.290	0.010
DE	2	410	205	0.804	0.003
DF	4	403	100	0.394	0.003
DG	4	260	65	0.254	0.002
EF	2	566	283	1.109	0.004
EG	2	163	81	0.319	0.001
FG	4	1,808	452	1.771	0.014
ABC	30	11,656	388	1.522	0.095
ABD	30	8,146	271	1.064	0.066
ABE	15	2,887	192	0.754	0.023
ABF	30	10,872	362	1.420	0.089
Residuals	3,646	930,485	255		

$$\hat{U}_i^{k,t}=\mu+\sum_{n=1}^{N}\phi_n\theta_n^k+\sum_{j=1}^{J}\sum_{\ell=1}^{L_j}\alpha_{j,\ell}\delta_{i,j,\ell}^{k,t}$$
$$+\sum_{j'=1}^{J}\sum_{\ell'=1}^{L_{j'}}\sum_{j=1}^{J}\sum_{\ell=1}^{L_j}\beta_{j,\ell,j',\ell'}\delta_{i,j,\ell}^{k,t}\delta_{i,j',\ell'}^{k,t}$$
$$+\sum_{j''=1}^{J}\sum_{\ell''=1}^{L_{j''}}\sum_{j'=1}^{J}\sum_{\ell'=1}^{L_{j'}}\sum_{j=1}^{J}\sum_{\ell=1}^{L_j}\gamma_{j,\ell,j',\ell',j'',\ell''}\delta_{i,j,\ell}^{k,t}\delta_{i,j',\ell'}^{k,t}$$
$$\times\delta_{i,j'',\ell''}^{k,t}+\ldots+\epsilon \quad (5)$$

where $\hat{U}_i^{k,t}$=estimated absolute route utility for route i, traveler k, and time t; μ=grand mean of estimated absolute route utilities; n=index of socio-demographic group, $n=1,\ldots,N$; ϕ_n=effect of the socio-demographic characteristics of group n; θ_n^k=binary index with value of 1 if traveler k belongs to group n, and 0 otherwise; j, j', j''=factor indices; ℓ, ℓ', ℓ''=level indices; L_j=number of levels of factor j; $\alpha_{j,\ell}$=effect of the ℓth level of factor j; $\delta_{i,j,\ell}^{k,t}$=binary index with value of 1 if level ℓ of factor j is perceived by traveler k along route i at time t, and 0 otherwise; $\beta_{j,\ell,j',\ell'}$=2nd level interaction or joint effect of the ℓth level of factor j, and the ℓ'th level of factor j'; $\gamma_{j,\ell,j',\ell',j'',\ell''}$=3rd level interaction effect of the ℓth level of factor j, the ℓ'th level of factor j', and the ℓ''th level of factor j''; and ϵ=independent random error term of normal distribution with zero mean and common variance.

Table 4. Route Utility Factorial Experiment Design Model Statistical Results

Source	Sum of squared errors	DF	Mean of squared errors	f-Statistic
Model	11,231,425	92	122,080	476
A	427,575	5	85,515	333
B	3,745,270	3	1,248,423	4874
C	963,169	2	481,584	1880
D	3,835,065	2	1,917,532	7486
E	768,831	1	768,831	3001
F	383,649	2	191,824	748
G	893,722	2	446,861	1744
AB	25,452	15	1696	6.6
AC	18,800	10	1880	7.3
AD	77,401	10	7740	30
AE	10,111	5	2022	7.8
AF	30,135	10	3013	11.7
AG	26,687	10	2668	10.4
BC	11,457	6	1909	7.4
BE	4,934	3	1644	6.4
CD	6,294	4	1573	6.1
CE	2,865	2	1432	5.5

Note: Model's R squared=0.920; Model's adjusted R squared=0.918.

In the FEDM, the following restrictions apply:

$$\sum_{n=1}^{N}\phi_n=0 \quad (6)$$

$$\sum_{\ell=1}^{L_j}\alpha_{j,\ell}=0 \quad ,\forall j=1,\ldots,J \quad (7)$$

$$\sum_{\ell'=1}^{L_{j'}}\sum_{\ell=1}^{L_j}\beta_{j,\ell,j',\ell'}=0 \quad ,\forall j=1,\ldots,J,\ \forall j'=1,\ldots,J \quad (8)$$

$$\sum_{\ell''=1}^{L_{j''}}\sum_{\ell'=1}^{L_{j'}}\sum_{\ell=1}^{L_j}\gamma_{j,\ell,j',\ell',j'',\ell''}=0 \quad ,\forall j=1,\ldots,J,$$
$$\forall j'=1,\ldots,J,\ \ j''=1,\ldots,J \quad (9)$$

It can be shown that under the above restrictions, there will be only one unique definition for each of the parameters in Eq. (5) (Miller et al. 1990). The parameters of the above model are determined using analysis of variance with the data collected from the score stage survey. The software *Design-Expert* (State-Ease 2000) was used for the FEDM calibration. In this paper, a model of only three-level interactions is considered.

Data Structure Preparation for Model Calibration

Table 2 shows a sample of the data structure for the FEDM calibration. Each data record (row) is comprised of seven categorical

Estimated Utility, $\hat{U}$ **=**

$$
\begin{aligned}
&+ 255.96 \quad \text{(Mean effect)} \\
&- 18.74 * A_1 + 13.02 * A_2 - 2.01 * A_3 - 4.71 * A_4 + 2.00 * A_5 \\
&+ 27.06 * B_1 + 24.15 * B_2 - 0.84 * B_3 \\
&+ 14.40 * C_1 + 7.50 * C_2 \\
&+ 33.66 * D_1 + 8.27 * D_2 \\
&- 14.06 * E \\
&- 12.94 * F_1 + 1.74 * F_2 \\
&- 19.65 * G_1 + 2.39 * G_2 \\
&+ 0.80 * A_1 B_1 + 1.09 * A_2 B_1 - 3.74 * A_3 B_1 + 2.85 * A_4 B_1 - 0.37 * A_5 B_1 \\
&+ 0.76 * A_1 B_2 + 2.48 * A_2 B_2 - 1.95 * A_3 B_2 + 1.63 * A_4 B_2 - 0.85 * A_5 B_2 \\
&- 0.72 * A_1 B_3 - 0.25 * A_2 B_3 - 0.17 * A_3 B_3 + 2.40 * A_4 B_3 - 1.19 * A_5 B_3 \\
&- 1.01 * A_1 C_1 + 2.61 * A_2 C_1 + 2.05 * A_3 C_1 - 3.48 * A_4 C_1 - 0.11 * A_5 C_1 \\
&+ 1.22 * A_1 C_2 + 0.88 * A_2 C_2 - 0.017 * A_3 C_2 - 2.50 * A_4 C_2 - 0.49 * A_5 C_2 \\
&- 8.20 * A_1 D_1 - 1.72 * A_2 D_1 + 0.32 * A_3 D_1 + 0.78 * A_4 D_1 + 5.06 * A_5 D_1 \\
&- 3.80 * A_1 D_2 - 0.90 * A_2 D_2 + 0.26 * A_3 D_2 + 1.83 * A_4 D_2 + 1.75 * A_5 D_2 \\
&+ 0.50 * A_1 E + 1.99 * A_2 E - 2.59 * A_3 E - 0.77 * A_4 E + 1.80 * A_5 E \\
&+ 2.97 * A_1 F_1 - 1.24 * A_2 F_1 - 2.86 * A_3 F_1 + 3.90 * A_4 F_1 + 3.46 * A_5 F_1 \\
&- 2.18 * A_1 F_2 + 1.22 * A_2 F_2 + 1.62 * A_3 F_2 - 0.089 * A_4 F_2 - 2.65 * A_5 F_2 \\
&- 1.16 * A_1 G_1 + 1.85 * A_2 G_1 + 5.53 * A_3 G_1 - 4.81 * A_4 G_1 - 2.16 * A_5 G_1 \\
&+ 0.95 * A_1 G_2 - 1.62 * A_2 G_2 - 0.79 * A_3 G_2 - 0.37 * A_4 G_2 + 0.88 * A_5 G_2 \\
&+ 4.17 * B_1 C_1 - 1.34 * B_2 C_1 - 1.66 * B_3 C_1 \\
&- 2.14 * B_1 C_2 + 0.98 * B_2 C_2 + 0.48 * B_3 C_2 \\
&- 1.92 * B_1 E + 0.42 * B_2 E + 0.97 * B_3 E \\
&+ 0.90 * C_1 D_1 - 0.098 * C_2 D_1 \\
&+ 1.50 * C_1 D_2 - 1.42 * C_2 D_2 \\
&- 0.94 * C_1 E - 0.20 * C_2 E
\end{aligned}
$$

(a) Route Utility FEDM

A: City area-Age Factor A_1: NC-Y : Non-congested area and young age (<25 years) A_2: NC-M : Non-congested area and middle age (25-45 years) A_3: NC-O : Non-congested area and old age (45+ years) A_4: C-Y : Congested area and young age (<25 years) A_5: C-M : Congested area and middle age (25-45 years) A_6: C-O : Congested area and old age (45+ years)		**B: Travel Time** B_1: Free B_2: Light B_3: Medium B_4: High
C: Highway Class C_1: Freeway C_2: Arterial with signal progression: AWP C_3: Arterial without signal progression: AWOP		**D: Queue Time** D_1: Light D_2: Medium D_3: Heavy
E: Familiarity E_1: Familiar E_2: Not-Familiar	**F: Pavement Condition** F_1: Poor F_2: Acceptable F_3: Good	**G: Speed** G_1: Low G_2: Medium G_3: High

(b) List of Abbreviations for Factors and Levels

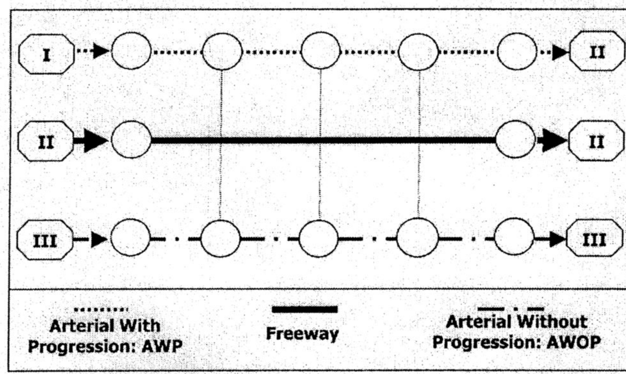
Fig. 3. Route utility factorial experimental design model

variables (A through G) and the average route utility. A total of 3888 records (with all possible factor-level combinations) were prepared. The A variable refers to the socio-demographic group, B variable refers to the travel time, C refers to the highway class, D refers to the queuing time, E refers to the familiarity, F refers to the pavement condition, and G refers to the speed. The last column represents the average utility of route i, $\overline{U^t_{i,g}}$ [estimated by Eq. (1)] for group g's participants. For example, the average utility in the first row (bold cell in Table 2) is calculated as follows:

$$\overline{U^t_{i,NC\text{-}Y}} = \frac{\sum_{k=1}^{K_{NC\text{-}Y}} \sum_{j=1}^{J} \sum_{\ell=1}^{L} S_j^k \hat{S}^k_{\ell,j} \delta^{k,t}_{i,j,\ell}}{K_{NC\text{-}Y}} \tag{10}$$

$\overline{U^t_{i,NC\text{-}Y}}$=average utility assigned by participants in the "NC-Y" group to route i at time t; and $K_{NC\text{-}Y}$=total number of participants in the "NC-Y" group.

The factor values (the first seven columns of Table 2) were coded. Coding refers to the transformation of the different measurement units (of factors) into the same common scale. Coding maps the factor ranges to a common scale, $[-1,1]$, regardless of their absolute (true) magnitudes. Factors' coding establish orthogonal levels so the factors' effects can be estimated independently (eliminate the correlation effect).



The following table summarizes the information of the above three routes. Please take your time reviewing the various route characteristics, and then indicate your favorite route below:

Factor	Type	Route Levels		
		I-I	**II-II**	**III-III**
Class	Categorical	AWP	Freeway	AWOP
Travel Time (minutes)	Numeric	10	12	13
Queue Time (minutes)	Numeric	4	2	6
Familiarity	Categorical	Familiar	Familiar	Familiar
Highway/Pavement Condition	Categorical	Poor	Acceptable	Good
Speed (km/hr)	Numeric	60	70	40
Decision: Please tick the appropriate box of your route selection		☐	☐	☐

Fig. 4. Sample template for validating surveying data and modeling

Analysis of Variance and FEDM Calibration

Table 3 summarizes the results of the analysis of variance. Three-level interaction effects were initially considered. The contribution (last column) represents the ratio of the squared errors attributed to the factor (or interaction) to the total squared errors. The travel time (B) contributes to 30.69% of the total squared errors. The queuing time (D) contributes to 31.4%. The highway class (C) and the speed (G) contribute to 7.9 and 7.3, respectively. The familiarity of the highway (E) contributes to 6.3. Finally, the group effect (A) and the pavement condition (F) contribute to 3.5 and 3.14, respectively.

Only the factors (and interactions) of significant f statistic (bold cells in Table 3) are included in the final model. All third-level interactions are statistically insignificant. All the second-level interactions with significant f statistic were included in the final form of the FEDM; even these with negligible contribution percentages (e.g., CE).

Table 4 shows the FEDM form and statistics. Slight changes are encountered in the f statistics (compared to initial results in Table 3) due to the exclusion of the insignificant effects. The *Model f* statistic of 476.65 implies that the model is significant with 0.01% probability that the *Model f* statistic is attributed to noise error. The R^2 of the derived factorial experiment model is about 0.92.

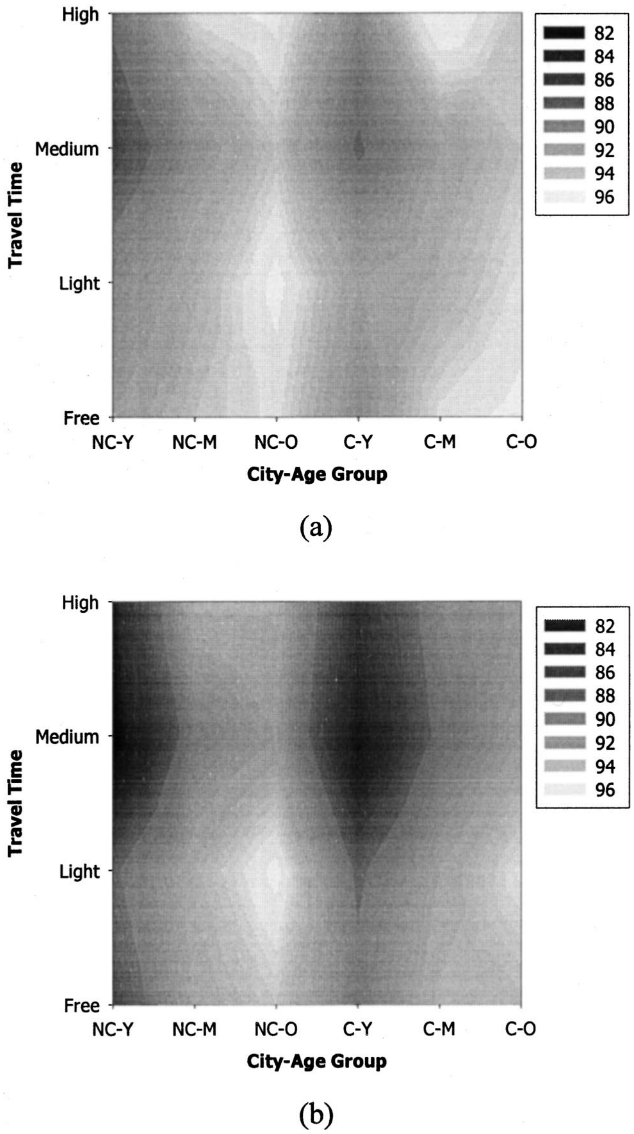


Fig. 5. Agreement percentages among (a) data estimated and actual decisions, (b) model estimated and actual decisions

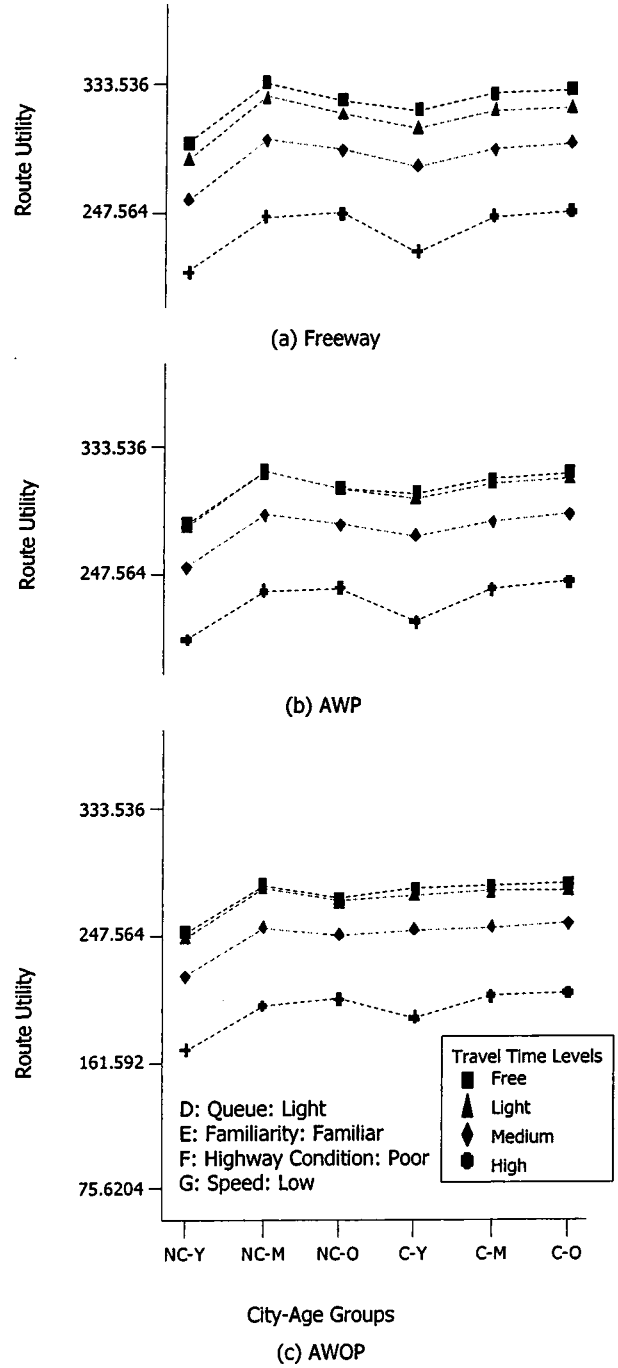


Fig. 6. Effect of highway class on estimated route utilities

Route Utility Factorial Experiment Design Model Estimation

Fig. 3 shows the final coded form of the route utility FEDM. It comprises three blocks; the mean effect, the factors' effects, and the selected second-level interactions' effects. For a two-level categorical factor (such as E), the coded coefficient is half the difference between the real coefficients of the first and second levels. As such, the E_1 level (Familiar) real coefficient would be 0 $(-14.06+14.06)$, and the E_2 (Nonfamiliar) coefficient would be -28.12 $(-14.06-14.06)$. For multilevel categorical factors (such as A), the coded coefficient (such as the -18.74 of A_1) is the difference between the mean effect and the utility perceived by a traveler from the A_1 group.

Similarly, the coefficient of B_1 represents the difference between the mean effect and the utility perceived for *free* travel time conditions. Finally, the coefficient of the last level of any factor (or interaction) is determined by the Eqs. (6) through (9). It is the negative sum of all the other levels' coded coefficients (e.g., $B_4=-(27.06+24.15-0.84)=-50.37$; $A_6B_1=-(0.80+1.09-3.74+2.85-0.37)=-0.63$. The estimated utility of any route is calculated as the sum of the mean effect and levels' real coefficients.

Modeling Approach Validation

The validation stage was conducted using PC-based experimental templates as shown in Fig. 4. A total of 500 templates were prepared to uniformly cover the various factors' levels (partial coverage). Each participant was assigned 20 randomly selected templates and asked to select among the three routes in the template. The actual decisions were then compared against the decisions estimated by the FEDM.

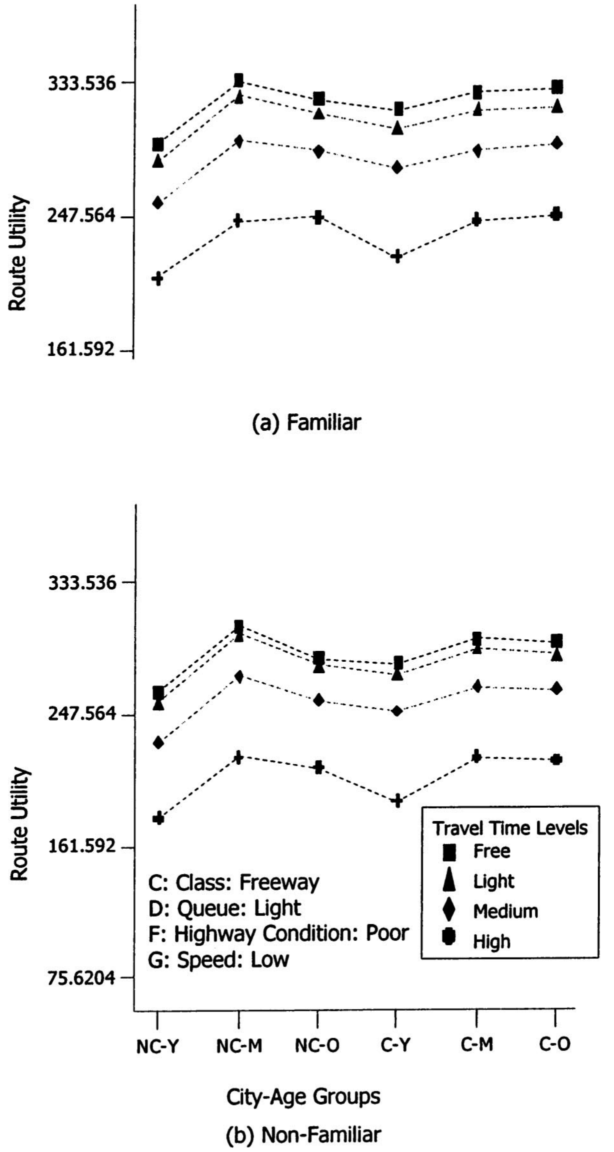


Fig. 7. Effect of route familiarity on estimated route utilities

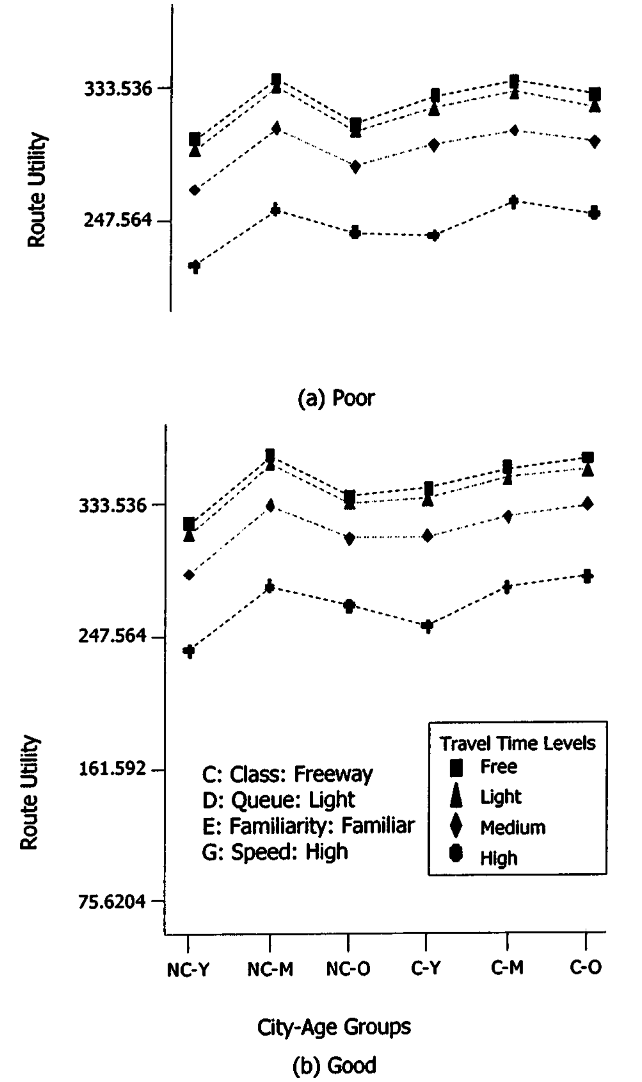


Fig. 8. Effect of highway/pavement condition on estimated route utilities

In brief, the validation is carried out in two stages:

1. Data validation: to assess the accuracy of the data from the score and numeric levels' identification stages and the validity of the scoring approach; and
2. Model validation: to assess the accuracy of the decisions estimated by the FEDM.

Data Validation Stage

The idea is to use the individual participant's scores and levels definitions to estimate the route utility (and hence the route choice) using the scoring technique. The participant's actual route is then compared with the estimated choice, and the discrepancy is assessed. The adopted procedure is summarized below.

1. The participant's record of the numeric level identification (Fig. 2) is retrieved. The values of the numeric factors (travel time, queuing time, and speed) in the validation templates (Fig. 4) are assigned linguistic terms (such as *high*, *low*, etc.) based on the participant numeric levels identification. For example, the speed of 40 km/h (for route III-III in Fig. 4) is assigned a *medium* term based on the participant identification of the speed terms (*low*: [0,30], *medium*: [30,60], *high*: [60, ∞[). Note that this step applies for the numeric factors only; the linguistic terms of the categorical factors (class, familiarity, and highway/pavement condition) are explicitly provided in the validation template.
2. The score record is also retrieved for the participant (Fig. 1). The linguistic terms (from step 1) are utilized to calculate the route utility using Eq. (1) for the three routes given the validation template.
3. The route with maximum utility is identified, and referred to as a data-estimated route.
4. The data-estimated route is compared against the actually selected route (in the validation template).
5. Steps 1 through 4 are repeated for the 20 templates randomly assigned to the participant.
6. Statistics are collected on the percentages of agreement among the data-estimated and corresponding actual routes.

Fig. 5(a) shows the percentages of agreement among the data-estimated and the actual decisions. The percentage ranges from 86 to about 97%. The darkest areas in the figure represent the highest discrepancy among data-estimated and actual decisions. The highest discrepancies are encountered for the NC-Y and C-Y groups (i.e., young-age participants). In addition, discrepancies are observed more often with the templates of high travel times. This suggests enriching the size of the young-age groups to increase the accuracy of the results. Generally, the results seems quite reasonable and within the expected range of accuracy, taking into consideration the time difference between the score and numeric level identification stages of data collection, and the time at which the validation procedure was carried out (nearly three weeks).

Model Validation Stage

The idea is to use the most probable factors' levels to estimate the route utility (and hence the route choice) using the developed FEDM. The participant's actual route is then compared with the FEDM-estimated choice, and the discrepancy is assessed. The adopted procedure herein is quite similar to that of the data validation stage. In validating the FEDM, the most probable levels for the group to which the participant belongs were utilized. For example, a template with travel time of 350% may be regarded differently by the participants of a specific group, but mostly as (say) a *medium* level. Therefore, the most probable term (*medium*) is used for estimating the *model* utility and the route choice decisions. Furthermore, instead of utilizing the factor and level scores (of the score stage), the FEDM (in Fig. 3) is utilized to estimate the route utility. This is referred to as model-estimated utility.

Fig. 5(b) illustrates the agreement percentages among the model-estimated and the actual decisions. The agreement percentages range from 82 to 95%. The maximum discrepancies are encountered for young-age groups, particularly with the templates illustrating high travel times. The discrepancies here may be attributed to the fitting errors, and the use of average group utility in the FEDM calibration.

Analyses of the Route Utility Factorial Experiment Design Model Results

The interactions among the factors significantly contribute to the value of the route utility. It might be misleading to assume (for example) that lesser utility is always associated with higher travel time. The significant factors' interactions may affect the utility values, and as such provide the necessity to study such interactions closely.

Fig. 6 summarizes the results for the participant groups. The route utility is given for the various travel-time levels. The figure illustrates the effect of the highway class for familiar routes. For a freeway, the route utility is clearly distinguished for the various travel time levels. For the arterial with progression (AWP) and the arterial without progression (AWOP), the route utilities are mostly identical for the free and the light levels of travel time. The route utility of C-Y group is among the least utilities for the freeway and the AWP, but among the highest for the AWOP routes. Note also that the difference in utility (for all groups) between the *free* and the *high* travel times is higher for the freeway, compared to the AWP and AWOP.

Fig. 7 illustrates the effect of the route familiarity. Note the close gap between the two curves of the *free* and the *light* travel times for a nonfamiliar freeway, compared to the familiar freeway.

Fig. 8 shows the effect of the highway and pavement condition. Note that the C-O group appreciates the *good* conditions more than the C-M group does (for all the travel time levels); the C-O group exhibited higher utilities. The reverse is truly correct for the *poor* highway conditions. Note also the difference between the curves of *medium* and *high* travel times for the NC-O and C-Y groups. At medium travel times, while the same utility is perceived for the good highway condition, the C-Y group perceives higher utility for the poor conditions.

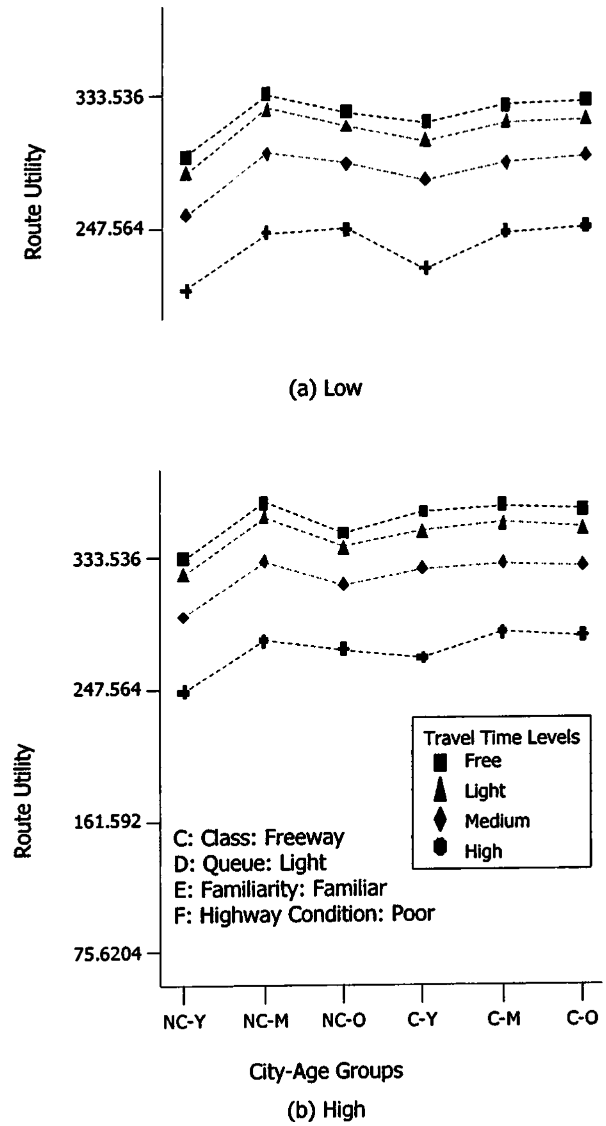


Fig. 9. Effect of speed on estimated route utilities

Fig. 9 illustrates the effect of the speed. While the utility is almost identical for all the age groups in congested city areas and high-speed level, the utility is reasonably less for the young age groups and low-speed conditions.

Fig. 10 shows the effect of the queuing time. Note in particular the difference between the *high* travel-time curves in the cases of *light* and *heavy* queuing time. While the utilities are almost identical among all the groups under heavy queuing times, it reasonably varies under light queuing times.

Concluding Comments

The limitations of the classical discrete choice models are twofold: the inability to capture the subjects' taste variation, and the vagueness of the subject behavior. With the recent research attempts, the taste variation was incorporated to some extent, yet

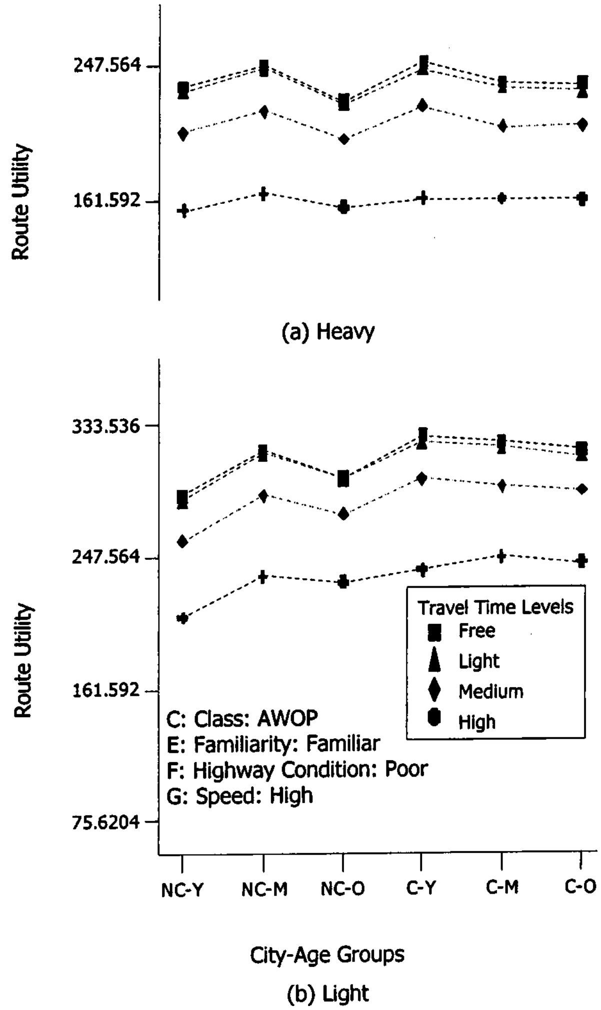


Fig. 10. Effect of queuing time on estimated route utilities

the vagueness still represents a restriction. This paper adopts the approach of stated preference rating of the route attributes for data collection. The FEDM may be utilized in assessing the significance of the various route attributes. It may also be utilized in studying the correlation and interaction effects among the various route attributes (rarely elaborated in the literature). The FEDM statistical results indicated that the route utility is highly influenced by the travel time, and queuing time (highest contribution percentages). Other factors such as the speed, highway classification, the highway condition, and the familiarity with the route are also among the important factors. The scoring technique was validated to ensure the accuracy of survey participants, and the validity of the developed factorial experimental design model. The interaction effects among the various factors at the different levels were also addressed.

It is imperative to realize that the FEDM could be used to capture the inter-group taste variation, but neither the intra-group taste variation nor the vagueness in the choice behavior. It is based on the assumption that the travelers within the same sociodemographic group will exhibit the same behavior if subjected to same attribute values. Future work will focus on developing the route choice models using stated preference ranking and choice data types.

Finally, it is worth noting that the main deficiency in using the FEDM is the unclear interpretations of the traffic states (levels) among the travelers; to calculate the route utility for a specific traveler, his/her own linguistic interpretations of the numeric traffic states are needed. The FEDM was calibrated using the *average* group scores, assuming the travelers of the same sociodemographic characteristics will perceive the same utility if subjected to the same numeric levels. Using the *average* interpretations of a specific city-age group might not be accurate. As such, this modeling approach cannot be utilized to capture the individual taste or the intra-group taste variation of utilities.

To overcome such limitations, a fuzzy logic is developed to model the route choice behavior. The fuzzy logic is trained via neural nets using the data collected during the numerical level identification stage, and the FEDM model-estimated utilities. Fuzzy logic can be used effectively in addressing the two issues of taste variation and vagueness, i.e., capturing the variability of the travelers' appraisal to the different route attributes, as well as the variability in their perceptions to the various attribute levels. The details of this logic are discussed in detail in (Hawas 2004).

Acknowledgments

This research has been funded by the Research Council at the UAE University under Grant Nos. 04-7-11/2000 and 05-7-11/20001. The writer expresses his sincere appreciation for all students, faculty members, and individuals who participated in the various survey stages indicated in this paper.

References

Abdel-Aty, M. A., Kitamura, R., and Jovanis, P. P. (1995a). "Exploring route choice behavior using geographic information system-based alternative routes and hypothetical travel time information input." *Transportation Research Record 1493*, Transportation Research Board, Washington, D.C., 74–80.

Abdel-Aty, M. A., Kitamura, R., and Jovanis, P. P. (1997). "Using stated preference data for studying the effect of advanced traffic information on drivers' route choice." *Transp. Res., Part C: Emerg. Technol.,* 5(1), 39–50.

Abdel-Aty, M. A., Kitamura, R., Jovanis, P. P., Reddy, P., and Vaughn, K. M. (1995b). "New approach to route choice data collection: Multiphase, computer-aided telephone interview panel surveys using geographic information systems data base." *Transportation Research Record 1493*, Transportation Research Board, Washington, D.C., 159–169.

Abdel-Aty, M. A., Vaughn, K. M., Kitamura, R., Jovanis, P. P., and Mannering, F. L. (1994). "Models of commuters' information use and route choice: Initial results based on southern California commuter route choice survey." *Transportation Research Record 1453*, Transportation Research Board, Washington, D.C., 46–55.

Adler, J. L. (1993). "A Conflict model and Interactive Simulator (FASTCARS) for predicting en-route driver behavior in response to real-time traffic condition information." *Transportation,* 20(2), 83–106.

Bonsall, P. W. (1992). "The Influence of route guidance advice on route choice in urban networks." *Transportation,* 19(1), 1–23.

Bonsall, P., Firmin, P., Anderson, M., Palmer, I., and Balmforth, P. (1997). "Validating the results of route choice simulator." *Transp. Res., Part C: Emerg. Technol.,* 5(6), 371–387.

Bonsall, P. W., and Parry, T. (1991). "Using an interactive route-choice simulator to investigate drivers' compliance with route guidance advice." *Transportation Research Record 1306*, Transportation Research Board, Washington, D.C., 59–68.

Chen, P. S., and Mahmassani, H. S. (1993). "Dynamic interactive simulator for studying commuter behavior under real-time traffic information supply strategies." *Transportation Research Record 1413*, Transportation Research Board, Washington, D.C., 12–21.

Emmerink, R. H. M., Nijkamp, P., Rietveld, P., and Van Ommeren, J. N. (1996). ''Variable message signs and radio traffic information: An integrated empirical analysis of drivers' route choice behavior.'' *Transp. Res. Rec.*, 30(2), 135–153.

Hawas, Y. E. (2004). ''Development and calibration of route choice utility models: Neuro-fuzzy approach.'' *J. Transp. Eng.*, 130(2), 171–182.

Henn, V. (2000). ''Fuzzy route choice model for traffic assignment.'' *Fuzzy Sets Syst.*, 116(1), 77–101.

Katsikopoulos, K. V., Duse-Anthony, Y., Fisher, D. L., and Duffy, S. A. (2000). ''The framing of drivers' route choices when travel time information is provided under varying degrees of cognitive load.'' *Hum. Factors*, 42(3), 470–481.

Kikuchi, S., and Pursula, M. (1998). ''Treatment of uncertainty in study of transportation: Fuzzy set theory and evidence theory.'' *J. Transp. Eng.*, 124(1), 1–8.

Koppleman, F. S., and Wen, C.-H. (2000). ''Paired combinatorial logit model: Properties, estimation, and application.'' *Transp. Res., Part B: Methodol.*, 34(2), 75–89.

Koutsopoulos, H. N., Lotan, T., and Yang, Q. (1994). ''Driving simulator and its application for modeling route choice in the presence of information.'' *Transp. Res., Part C: Emerg. Technol.*, 2(2), 91–107.

Koutsopoulos, H. N., Polydoropoulou, A., and Ben-Akiva, M. (1995). ''Travel simulators for data collection on driver behavior in the presence of information.'' *Transp. Res., Part C: Emerg. Technol.*, 3(3), 143–159.

Kuri, J., and Pursula, M. (2001). ''Modeling car route choice with nondynamic stated preference data: Comparison between fuzzy similarity and random utility models.'' *CD-ROM Proc., 80th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Lotan, T. (1997). ''Effects of familiarity on route choice behavior in the presence of information.'' *Transp. Res., Part C: Emerg. Technol.*, 5(3), 225–243.

Lotan, T. (1998). ''Modeling discrete choice behavior based on explicit information.'' *IEEE Trans. Syst. Man Cybern.*, 28(1), 100–114.

Mahmassani, H. S., and Herman, R. (1990). *Interactive experiments for the study of trip-maker behavior dynamics in congested commuting systems. Developments in dynamic and activity-based approaches to travel analysis*, P. Jones, ed., Gower, Aldershot, Hampshire, U.K., 272–298.

Miller, I., Freund, J. E., and Johnson, R. A. (1990). *Probability and statistics for engineers*, 4th Ed., Prentice-Hall, Englewood Cliffs, N.J.

Ozbay, K., Datta, A., and Kachroo, P. (2001). ''Modeling route choice behavior using stochastic learning automata.'' *CD-ROM Proc., 80th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Papacostas, C. S., and Prevedouros, P. D. (1993). *Transportation engineering and planning*, 2nd Ed., Prentice-Hall, Englewood Cliffs, N.J.

Peeta, S., Ramos, J. L., and Pasupathy, R. (2000). ''Content of variable message signs and on-line driver behavior.'' *Transportation Research Record 1725*, Transportation Research Board, Washington, D.C., 102–108.

Polydoropoulu, A., Ben-Akiva, M., and Kaysi, E. (1994). ''Influence of traffic information on drivers' route choice behavior.'' *Transportation Research Record 1453*, Transportation Research Board, Washington, D.C., 56–65.

Rilett, L. R., and Park, D. (2001). ''Incorporating uncertainty and multiple objectives in real-time route selection.'' *J. Transp. Eng.*, 127(6), 531–539.

Sermons, M. W., and Koppleman, F. S. (1998). ''Factor analytic approach to incorporating systematic taste variation into models of residential location choice.'' *Transportation Research Record 1617*, Transportation Research Board, Washington, D.C., 194–202.

Srinvasan, K. K., and Mahmassani, H. S. (2000). ''Modeling inertia and compliance mechanisms in route choice behavior under real-time information.'' *Transportation Research Record 1725*, Transportation Research Board, Washington, D.C., 45–53.

State-Ease. (2000). *Design expert software: Version 6 user's guide*, State-Ease, Inc.

Swait, J., and Bernardino, A. (2000). ''Distinguishing taste variation from error structure in discrete choice data.'' *Transp. Res., Part B: Methodol.*, 34(1), 1–15.

Teodorovic, D., Vukanovic, S., and Obradovic, K. (1998). ''Modeling route choice with advanced traveler information by fuzzy logic.'' *Transp. Plan. Technol.*, 22(1), 1–25.

Vaughn, K. M., Abdel-Aty, M. A., Kitamura, R., Jovanis, P. P., Yang, H., Kroll, N. E. A., Post, R. B., and Oppy, B. (1993). ''Experimental analysis and modeling of sequential route choice under an advanced traveler information system in a simplistic traffic network.'' *Transportation Research Record 1408*, Transportation Research Board, Washington, D.C., 75–82.

Vaughn, K. M., Kitamura, R., and Jovanis, P. P. (1995b). ''Experimental analysis and modeling of advice compliance: Results from advanced traveler information system simulation experiments.'' *Transportation Research Record 1485*, Transportation Research Board, Washington, D.C., 18–26.

Vaughn, K. M., Reddy, P., Abdel-Aty, M. A., Kitamura, R., and Jovanis, P. P. (1995a). ''Route choice and information use: Initial results from simulation experiments.'' *Transportation Research Record 1516*, Transportation Research Board, Washington, D.C., 61–69.