Transportation (2006) 33: 63–82
DOI: 10.1007/s11116-005-5354-y
© Springer 2006

Modeling travel time under ATIS using mixed linear models

M. FATHY ABDALLA & MOHAMED ABDEL-ATY*
Department of Civil & Environmental Engineering, University of Central Florida,
Orlando, FL, 32816-2450, USA
*(*Author for correspondence, E-mail: mabdel@mail.ucf.edu)*

Key words: ATIS, link choice, mixed linear models, repeated observations, travel time

Abstract. The objective of this paper is to model travel time when drivers are equipped with pre-trip and/or en-route real-time traffic information/advice. A travel simulator with a realistic network and real historical congestion levels was used as a data collection tool. The network included 40 links and 25 nodes. This paper presents models of the origin-to-destination travel time and en-route short-term route (link) travel time under five different types and levels of advanced traveler information systems (ATIS). Mixed linear models with the repeated observation's technique were used in both models. Different covariance structures (including the independent case) were developed and compared. The effect of correlation was found significant in both models. The trip travel time analysis showed that as the level of information increases (adding en-route to the pre-trip and advice to the advice-free information), the average travel time decreases. The model estimates show that providing pre-trip and en-route traffic information with advice could result in significant savings in the overall travel time. The en-route short-term (link) travel time analysis showed that the en-route short-term (link) information has a good chance of being used and followed. The short-term qualitative information is more likely to be used than quantitative information. Learning and being familiar with the system that provides the information decreases en-route short-term delay.

1. Introduction

Advanced traveler information systems (ATIS) are intended to assist travelers in planning and decision making for mode, departure time, and route choices, including congestion avoidance, to improve the convenience and efficiency of travel. The impact and effectiveness of ATIS, however, critically depend on traveler's responses to these systems and to the information that they offer. Therefore, it is essential to understand the traveler's decision-making process under real-time information. The objective of this paper is to model travel time and en-route short-term link travel time when drivers are equipped with pre-trip and/or en-route real-time traffic information/advice. A travel simulator was used to collect dynamic route choice data. The simulator uses realistic network and real historical volumes. Different weather conditions are also used. The simulator accounts for delays caused by intersections, recurring congestion, non-recurring congestion (incident), and queuing at toll plazas. The network consists of 25 nodes and 40 links and comprises different types of highways. For a detailed design and description of the simulator and the network, the reader is referred to Abdalla (2003). Abdel-Aty and Abdalla (2002) illustrated the validity of the data and the simulator as a route-choice data collection tool. Figure 2 shows a spot view from the simulator.

The problem of repeated measurements arose in this study because each subject made multiple choices. These choices are correlated. This correlation must be taken into account. Otherwise, the model would underestimate (overestimate) the standard errors of the between- (within-) subject effects (Stokes et al. 2000). Mixed linear models with different covariance structures were used and compared with the independent case to validate the statistical analysis.

2. Background

2.1. *Previous methodologies for modeling driver's behavior under ATIS*

Researchers have used cross tabulation, analysis of variance, and probabilistic models in studying and modeling driver's behavior under ATIS (Vaughn et al. 1995a, b; Koutsopoulos et al. 2000). Driver behavior modeling has been traditionally modeled by maximizing a certain utility function; mostly logit or probit models to understand driver behavior under ATIS. Abdel-Aty et al. (1994a) estimated three models: (1) a bivariate probit model of whether commuters access pre-trip information and whether they use multiple routes, (2) a bivariate probit model of whether commuters access en-route information and use multiple routes, and (3) a negative binomial model of the frequency of route changes given pre-trip or en-route information use. Abdel-Aty et al. (1994b) developed a binary logit model to estimate respondent's choice to accept or reject an ATIS advice. Abdel-Aty et al. (1995) developed a model that used five hypothetical binary choice sets collected in a computer-aided telephone interview to determine how travel time variation affects route choice, and the potential interplay among travel time variation, traffic information acquisition and route choice. Khattak et al. (1995) estimated a bivariate ordinal probit model of driver's willingness to change route or departure time given traffic information. They used two dependent variables with five-point scale responses indicating the degree of driver's agreement (“strongly agree” to “strongly disagree”) that they would change aspects of their travel. Liu and Mahmassani (1998) used a multivariate probit model with 24 dependent variables that were 4 days worth of decisions to change departure time, route before leaving home, and route at each of four en-route decision points. This model took into account the commuter's learning from experience. Chen et al. (1999) employed event-count (frequency) models to capture the principal effects of the commuter's experience with real-time information on user compliance. Yamamoto et al. (2000) developed a multinomial logit model with alternatives representing the choice between freeways and surface streets.

2.2. *Previous driver's behavior modeling with repeated observations*

When the same respondent makes multiple choices, the dependent variable and then the error terms for these choices are correlated. Few transportation-related efforts have been conducted to account for this correlation. Louviere and Woodworth (1983) and Mannering (1987) corrected the standard errors produced by a repeated responses regression model by simply multiplying the standard errors by the square root of the number of repeated observations. Kitamura and Bunch (1990) used a dynamic ordered-response probit model of car ownership with error components. Mannering et al. (1994) used an ordered logit probability model and a duration model with heterogeneity correlation term. Morikawa (1994) used logit models with error components to treat serial correlation. Abdel-Aty et al. (1997) addressed this issue using individual-specific random error components in binary logit models with a normal mixing distribution. The standard deviation of the error components were found to be significant. This showed clearly the need for some formal statistical corrections to account for the unobserved heterogeneity. Jou and Mahmassani (1998) used a general probit model form for a dynamic switching model, allowing the introduction of state dependence and serial correlation in the model specification. Mahmassani and Liu (1999) used a multinomial probit model framework to capture the serial correlation arising from repeated decisions made by the same respondent. Garrido and Mahmassani (2000) used a multinomial probit model with spatial and temporally correlated error structure. Chen and Jovanis (2003) used a mixed linear model with repeated observations to model driver's compliance with en-route guidance.

2.3. *Travel time saving due to ATIS*

A considerable number of studies have examined the potential benefits of providing pre-trip and en-route real-time information to travelers. Researchers are interested in the effects of ATIS on all types of travel decisions. ATIS is empirically shown to result in lessening travel time, congestion delays, and incident clearance time (Wunderlich 1996; Abdel-Aty et al. 1997; Sengupta & Hongola 1998). There is empirical evidence supporting the hypothesis that travelers alter their behavior in response to ATIS (Bonsall & Parry 1991; Vaughn et al. 1995a, b; Zhao et al. 1996; Mahmassani & Hu 1997). Reiss et al. (1991) have reported travel time savings that ranged from 3% to 30% and reduction in incident and congestion delays of up to 80% for impacted vehicles. However, other studies argued that providing information might not necessarily reduce congestion (Arnott et al. 1990).

Based on the above review, there is a need for further understanding of the decision process underlying traveler's behavior in the presence of ATIS, and in particular the implications for travel time. The literature is in need of a study that groups most of the previously investigated factors together. The analysis of this paper includes driver's socioeconomics, driving experience, driver's familiarity with pre-trip/en-route traffic information, the existence of five different levels of ATIS, different weather conditions, familiarity with the network and familiarity with the device that provides the information (the learning effect).

Moreover, most related analyses (with few exceptions, some of which are mentioned above) ignored the correlation between repeated decisions made by the same traveler. It has also been concluded that the literature needs more efficient and statistically approved methodologies to handle this problem, which may bias the results. Gopinath (1995) demonstrated that different model forecasts result when heterogeneity of travelers is not considered. Delvert (1997) argued that models of travel behavior in response to ATIS must address heterogeneity in behavior. To draw accurate conclusions from repeated-choices data, an appropriate model of within-subject correlation must be used. If correlation is ignored by using a model that is too simple, the model would underestimate the standard errors of the within-subject effects, and overestimate the standard errors of the between-subject effects (Stokes et al. 2000). On the other hand, if too complex a model is used, the analyst loses power and efficiency. In this paper, unobserved heterogeneity is considered by specifying and incorporating three different covariance structures of the correlated choices in the modeling process.

3. Methodology

3.1. *Mixed linear models*

A mixed linear model is a generalization of the standard linear model. The generalization being that the data is permitted to exhibit correlation and non-constant variability. The mixed linear model, therefore, provides the flexibility of modeling not only the means of the data (as in the standard linear model) but their variances and covariances as well. The standard linear model is certainly a useful one (Searle 1971). However, the distributional assumption about its error term vector ε is too restrictive (SAS 2003). In the mixed linear models, two different types of effects can be included, fixed-effects and random-effects. The fixed-effects are similar to those in the standard linear models. A random-effect is a variable that clusters the data where within-cluster correlation exists. In fact, the combination of these two types of effects led to the name mixed model. The reader is referred to Searle et al. (1992) for historical developments of the mixed models. The SAS/STAT PROC MIXED is a useful application for modeling mixed linear models with repeated observations.

3.2. *Modeling correlation in mixed linear models*

The mixed model extends the standard linear model by allowing a more flexible specification of the covariance matrix of ε. In other words, it allows for both correlation and heterogeneous variances. The mixed model is written as

$$y = X\beta + Z\gamma + \varepsilon$$

where y is the vector of the response variable, X is the known matrix of the explanatory effects, β is unknown fixed-effects parameter vector, Z is the known matrix of the random effects, γ is unknown random-effects parameter vector, ε is the vector of measurement errors (assumed independent and identically distributed (IID) in the case of standard linear model), and γ and ε are normally distributed with covariance matrices G and Σ, respectively. The variance-covariance matrix of Y is $V=(ZGZ'+\Sigma)$.

Likelihood-based methods are used to get β, SAS PROC MIXED employs two different methods: maximum likelihood (ML) and restricted maximum likelihood (REML). REML was favored, in this paper, because its estimators are obtained not from maximizing the whole likelihood function, as in the case of ML, but only the part that is invariant to the fixed-effect parameters vector β (Littell et al. 1996).

3.3. *Covariance structure V of the error components ε*

Three different variance-covariance structures (forms of Σ; recall, in this paper $V=\Sigma$) are used including the independent case. Verbeke and Molenberghs (2000) provide more forms for Σ. Guerin and Stroup (2000) documented the effects of various variance-covariance structures.

1. Independent Structure, it assumes complete independence over the error term components. This means that any two observations in the data

are independent from each other. It has one known variable (σ^2) that is the variance of the residuals.

$$\mathrm{cov}(y_{ij}, y_{ik}) = \begin{cases} \sigma^2 & \mathrm{j=k} \\ 0 & \mathrm{j \neq k} \end{cases} \quad \text{e.g.,} \rightarrow \Sigma_{3\times3} = \begin{bmatrix} \sigma^2 & 0 & 0 \\ & \sigma^2 & 0 \\ & & \sigma^2 \end{bmatrix} = \sigma^2 I_{3\times3}$$

2. Compound symmetry (CS), it assumes constant variance and constant covariance for the error terms. This means that the correlation between any two repeated choices is equal. Therefore, the error components for all the observations are equal. It has 2 unknown variables to be estimated by REML.

$$\mathrm{cov}(y_{ij}, y_{ik}) = \begin{cases} \sigma^2 + \sigma_1^2 & \mathrm{j=k} \\ \sigma_1^2 & \mathrm{j \neq k} \end{cases} \quad \text{e.g.,} \rightarrow \Sigma_{3\times3} = \begin{bmatrix} \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ & \sigma^2 + \sigma_1^2 & \sigma_1^2 \\ & & \sigma^2 + \sigma_1^2 \end{bmatrix}$$

3. Unstructured (UN), it specifies a completely general covariance matrix parameterized directly in terms of variances and covariances. It assumes different correlation between any two choices within a subject. This means that the correlation between any two repeated choices made by a certain subject is independent of the correlation between any other two choices made by the same subject. Therefore, the error component for each observation is independent from the other error components. It has $(n_i(n_i+1)/2)$ unknown parameters to be estimated by REML.

$$\mathrm{Corr}(y_{ij}, y_{ik}) = \sigma_{ij} \quad \text{e.g.,} \rightarrow \Sigma_{3\times3} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ & \sigma_2^2 & \sigma_{23} \\ & & \sigma_3^2 \end{bmatrix}$$

4. Simulator description

The Orlando transportation experimental simulation program (OTESP) was used to collect multidimensional route choice data. OTESP is an interactive windows-based computer simulation tool. It simulates commuter home-to-work morning trips. A portion of the city of Orlando network was captured from a GIS database (Figure 1). The network is located in an urban area and consists in 25 nodes and 40 links. This network has been carefully chosen from the Orlando network. It comprises different types of highways. It includes a 6-lane principle arterial, a 4-lane principle arterial, a 6-lane minor arterial, a 2-lane minor arterial, and local collectors. The network also includes two expressways.

The subject is presented with a real map of the network and has the ability to move his/her vehicle on the network from one intersection to another.

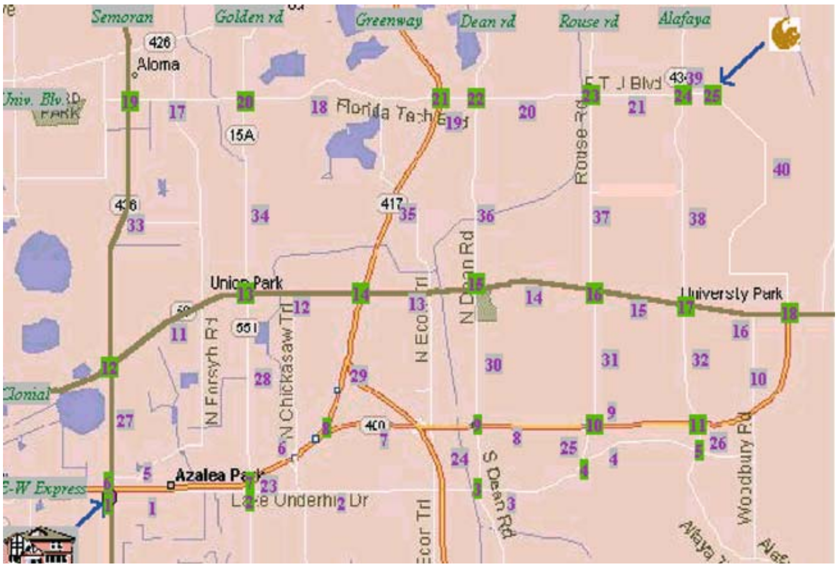

Figure 1. OTESP's network: codes for its nodes and links.

OTESP's code is fed with real historical traffic volumes on all network links. Based on these volumes, the geometric characteristics of the network roadways, and the HCM standards, the simulator generates a random speed for every link on the network. This generated speed accounts for delays caused by intersections, recurring congestion, non-recurring congestion (incident), queuing at toll plazas, and weather condition effects. All link's speeds are updated every time the subjects approach an intersection. The generated speeds on connected links are correlated because they are derived from real volumes which are correlated. The speed of a link, at a certain movement, controls the speed of the simulated vehicle so that the subject feels the delay. OTESP was distinctively designed to give the subject the feel of a realistic situation. Therefore, the simulator does not interfere with the accuracy of the information, whatever speed is generated for a particular movement (whether the travel time is provided to the subject or not) is assumed accurate for the current movement. The simulator provides the travel time value on every link as quantitative information while every link on the map takes a color representing its traffic congestion (green, yellow, or red for free, moderate, or congested flow). Therefore, the information provided is both descriptive and prescriptive. Figure 2 shows an example of the OTESP's user-interface and network.

At every trial day, OTESP provides one of five different scenarios (levels) of information/advice to the subject including: no information (scenario #1), pre-trip information without and with advice (scenarios #2 and #3, respec-

Figure 2. A spot view for OTESP (shown is scenario #5).

tively), and en-route information (in addition to the pre-trip information) without and with advice (scenarios #4 and #5, respectively). The subject is presented with these 5 scenarios, respectively, i.e., Scenario 1 then 2 until 5. Then the simulator repeats these 5 scenarios again to the subject and in the same order. This order facilitates the analysis process and the comparisons between the scenarios. During the actual experiment, the subject is required to complete these ten simulated days (two days for each scenario). In the 5 scenarios, the information (pre-trip en-route) is provided in two forms. First, giving all links on the network colors that represent their congestion levels (green, yellow, or red for free, moderate, or congested flows, respectively). Second, the travel time of every link is given. These two forms represent qualitative and quantitative information, respectively. Scenario 5 also provides an advised route from the subject's location to the destination.

The first 5 trial days for every subject (one per scenario) are named the first-trial-days. Similarly, the next 5 trial days are named the last-trial-days. In the first-trial-days, the subjects were assumed to be non-familiar with the device that disseminates the information. On the other hand, they were assumed familiar in the last-trial-days. There is no difference in the travel time computations across the scenarios. The differences between the five scenarios are only in the level of the information/advice provided to the subjects and whether they are pre-trip or en-route. OTESP also provides three different weather conditions (clear sky, light rain, and heavy rain). The Moore's shortest path algorithm (Pallottino & Grazia 1988) has been employed in the code of OTESP to determine the travel-time-based shortest path from any node to the destination, which is introduced as advice to the subjects in some scenarios. The simulator starts and ends with a short survey to collect the subject's socio-demographic characteristics, preferences, perceptions, and feedback. A four-table database is created to capture all the information/advice provided and the traveler decisions. For detailed design and description of the simulator and the network, the reader is referred to Abdalla (2003) and Abdel-Aty and Abdalla (2002).

5. Data collection

Subjects were recruited from the University of Central Florida (UCF) using an experimental design based on their age and gender. They included faculty, staff, graduate and undergraduate students. In this paper, the driver characteristics that were investigated include gender, age, income, level of education, and driving experience. Students were chosen to represent low-income, young, less driving experience, and moderate level of education. Faculty members were chosen to represent high level of education and income. Staff members were chosen to fill up the empty cells in the experimental design as they include a wide range of age and income. In addition, the authors meant to recruit all subjects from UCF (which is the destination of the network used) to be their real destination in their morning commute trips. About 12% of the subjects were in their first 2 weeks of attending UCF, representing non-familiar travelers. While we acknowledge that the sample is not a random sample of the population due to funding limitations, as mentioned above a concerted effort was carried out to obtain a representative sample based on an experimental design from UCF (40,000 student population + 2000 faculty and staff).

Subjects were instructed that their main task is to minimize the overall trip travel time by deciding when to and when not to follow the information and/or advice provided. Subjects have been asked not to go through the simulation unless they had at least 30 mins of spare time (the average simulation time was found to be 23.77 mins) and were willing to concentrate and do their best in their choices. Moreover, during the simulation, the subject's response time was measured without notifying them, to insure that they were serious. A total of 65 subjects had run the simulation. Two subjects out of the 65 have been excluded from this study because their response time were found to be outliers in the normal distribution plotting of subject's response time ($Z = 3.21$ and 3.78, $Z_{cr} = 2.57$).

6. Travel time analysis

A total of 630 trial days (trips) were completed by the 63 qualified subjects. Out of these 630 trial days, 539 were in the drive mode and 91 were in the transit (bus) mode. The analysis of this paper focuses on the drive mode. The simulator calculates and saves the cumulative total travel time of the trip as the subject moves the simulated vehicle until the destination. The overall average travel time of the 539 trial days was found to be 21.7 mins. The results showed that scenario 1 held the maximum average travel time (26.45 mins). Figure 3 provides a graphical display for the average travel time for the first and last trial-days of each scenario. Average travel time in all scenarios was less in the last trial-days indicating a learning effect which leads to drivers making route decisions that reduce their travel time. Figure 3 shows that the average travel time gradually decreases from scenario 1 to 5. This means that as the level of information/advice increases the travel time decreases.

6.1. *Modeling the trip travel time*

The aforementioned analysis shows the positive influence of providing traffic information/advice on the travel time. A mixed linear model with repeated observations was used to get better understanding of this influence and to

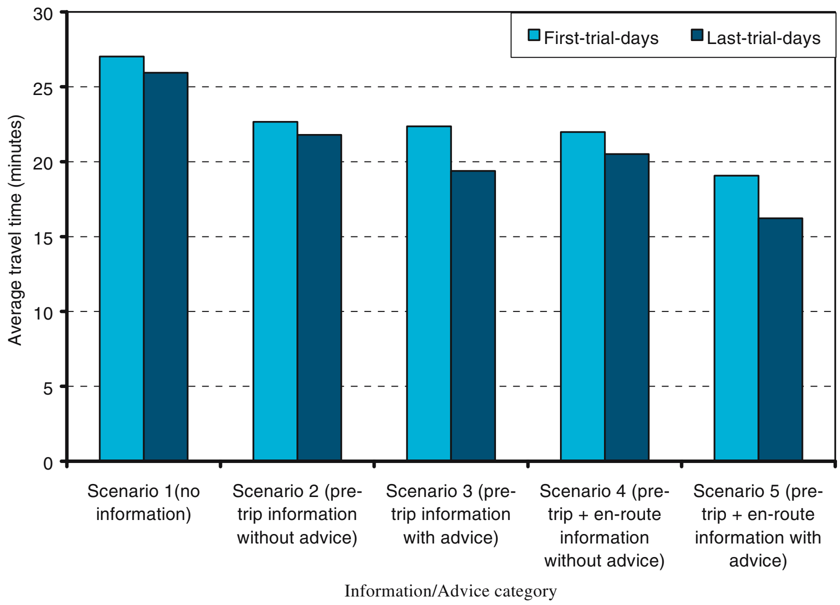

Figure 3. Average travel time versus information/advice category.

measure the marginal effects of the different significant variables that affect the travel time. Three different variance-covariance structures were used and compared (Independent case, CS, and the UN) to ensure the validity of the statistical analysis. The 539 trips were in the data used. The travel time is the dependent variable. The explanatory variables are:

1. Education level; "1" if graduate school or higher, "0" otherwise
2. Information familiarity; "1" if subject uses pre-trip and/or en-route traffic information usually or everyday, "0" otherwise
3. Network familiarity; "1" if subject is familiar with the network, "0" otherwise
4. Pre-trip information; "1" for scenario 2, "0" otherwise
5. Pre-trip information/advice; "1" for scenario 3, "0" otherwise
6. En-route information; "1" for scenario 4, "0" otherwise
7. En-route information/advice; "1" for scenario 5, "0" otherwise
8. System learning; "1" for the last-trial-days of the experiment, "0" for the first-trial-days
9. Light rain; "1" for light rain condition, "0" for heavy rain or clear sky
10. Heavy rain; "1" for heavy rain condition, "0" for light rain or clear sky

Table 1 shows the parameter estimates with a comparison between the three covariance structures CS, and UN as well as the independent. The maximum number of repeated choices per subject was ten. Therefore, the number of unknown and estimated parameters in the covariance matrices were [1], [2], [10 (11)/2 = 55] for the Independent case, CS, HF, and UN, respectively. Shown in Table 1, the likelihood ratio test (between the convergent versus restricted models) for the three structures proved that the explanatory variables could statistically explain the response variable. These three structures are nested within each other, or in other words, one is a special case of the other (Littell et al. 1996). Then a restricted likelihood ratio test can be performed with the test statistic equal to twice the difference between the two log likelihoods which follows a Chi-squared distribution with degrees of freedom that equals the difference in the number of parameters to be estimated in the two covariance matrices. This test can be used to favor one structure over the other (SAS, 2003). For example, to compare between the CS and UN structures, the χ^2 test static was 124.6 with degrees of freedom $= 53$ (55 −2). Therefore, the test was significant and favored the UN structure over the CS structure. The bottom part of Table 1 has the results of three similar Chi-square tests. These 3 tests proved that the UN structure is favored over the other two structures. Akaike information criteria (AIC), better model has a smaller AIC value (Wolfinger & Chang 1996), agreed with these results as the UN structure has the least AIC value. Therefore, the unstructured covariance matrix is the best structure for this model. This

Table 1. Results of long-term travel time analysis.

	Independent		CS		UN	
	Coeff.	t-static	Coeff.	t-static	Coeff.	t-static
Intercept	25.45	26.10	25.48	25.17	26.40	28.99
Education; 1 if graduate school or higher, 0 otherwise	1.60	2.75	1.56	2.49	1.38	2.70
Info. familiarity; 1 if subject uses pre-trip and/or en-route traffic information usually or everyday, "0" otherwise	−4.03	−2.11	−4.02	−1.93	−2.82	−1.88
Network familiarity; 1 if subject is familiar with the network, "0" otherwise	−1.02	−1.51	−1.00	−1.38	−1.29	−2.14
Pre-trip information; 1 for scenario 2, "0" otherwise	−4.63	−5.71	−4.64	−5.78	−4.50	−5.87
Pre-trip information/advice; 1 for scenario 3, "0" otherwise	−6.02	−7.56	−6.02	−7.63	−6.20	−7.62
En-route information; 1 for scenario 4, "0" otherwise	−5.57	−6.88	−5.58	−6.95	−5.50	−6.93
En-route information/advice; 1 for scenario 5, "0" otherwise	−9.07	−11.39	−9.09	−11.52	−9.72	−14.56
System learning; 1 for the last-trial-days of the experiment, "0" for the first-trial-days	−1.93	−3.76	−1.95	−3.83	−2.39	−5.38
Light rain; 1 for light rain condition, 0 for heavy rain or clear sky	2.48	4.13	2.44	4.08	1.98	4.47
Heavy rain; 1 for heavy rain condition, 0 for light rain or clear sky	4.08	6.15	4.10	6.18	3.69	8.24
Summary statistics						
Sample size = 539						
Log likelihood at zero [L (0)]	−1806.3		−1806.2		−1747.7	
Log likelihood at convergence [L (β)]	−1714.3		−1713.9		−1651.6	
LR statistic = 2 [L(β)−L (0)], D.O.F = 10	184.1		184.7		192.3	
AIC	3430.5		3431.7		3413.1	

Summary of chi-square test results to compare between different covariance structure:
LR (CS versus Indep.); $\chi^2_{\text{stat.}=0.8}$; D.O.F = 2–1 = 1; Not significant.
LR (CS versus UN); $\chi^2_{\text{stat.}=124.6}$; D.O.F = 55–2 = 53; Favors UN.
LR (UN versus Indep.); $\chi^2_{\text{stat.}=125.4}$; D.O.F = 55–1 = 54; Favors UN.

also shows that the correlation between repeated choices made by the same subject is significant in this model and had to be accounted for.

Table 1 shows the coefficient estimates. The estimate of the intercept was relatively high when compared to the other coefficients. This was reasonable, because the intercept represents the expected average travel time when all explanatory variables are zero. The model results showed that highly-educated drivers had longer travel times. Drivers who were familiar with traffic information and those who were familiar with the network had relatively less travel time. Providing pre-trip information without advice could result in a relative reduction in travel time of 4.5 mins compared to no information (4.5/21.71 = 20.7% of the average origin-to-destination travel time on the network). Providing pre-trip information with advice reduces the travel time by 28.6%. Providing pre-trip and en-route information without advice reduces the travel time by 25.3%. Providing pre-trip and en-route information with advice provide relative reduction of the travel time by 44.7%. By comparing the last four variables, it can be noticed that adding advice to the advice-free pre-trip information increases the travel time saving from 20.7% to 28.6%. Similarly, adding advice to the advice-free en-route information increases the travel time saving from 25.3% to 44.7%. Adding en-route to the pre-trip information increases the travel time saving from 20.7% to 25.3% in the case of providing information without advice and from 28.6% to 44.7% in the case of information with advice. In general, the travel time saving ranges from 20.7% to 44.7% due to ATIS. Many studies have looked into travel time saving due to ATIS, however, only a few of them have looked into and compared the travel time saving under more than one type/level of ATIS. This model analyzes the origin-to-destination travel time under 5 different types and levels of ATIS. The results also showed that familiarity with the system that provides the information reduces the travel time by 11.01%. Light and heavy rain conditions increase the travel time by 9.12% and 17%, respectively. Other factors including age, gender, income, and driving experience were tested and found not correlated with the travel time saving due to ATIS.

Table 2 shows the estimated 10-by-10 symmetric unstructured correlation matrix. The maximum correlation was 0.473 and the minimum was 0.001. We were not able to derive any evidence of a fixed trend of how the correlation between the choices increases (decreases) depending on the order of the choices. This underscores an open point for future research.

7. Short-term travel time (link) analysis

At each node in the simulator, the subject is required to make a decision and choose between the coming links. It was noticed that all the link movements had the direction west-to-east or south-to-north (this was expected because the destination is located north-east of the origin). Therefore, it was assumed that there are two link-alternatives (the east and north links) when the subject reaches a new node. The simulator assists the subject to choose between these two links by providing short-term (link) traffic information in scenarios #4 and 5 (scenarios #1, 2 and 3 provide no en-route traffic information).

Throughout this study, 4753 movements (decisions) have been made through the 40 links of the network. Out of them, 1667 movements were excluded from the analysis because the subject had no choice but to proceed on to the unique coming link. This happens at the most east, west, and north nodes of the network (see Figure 1). Out of the remaining 3086, Figure 4 shows the percentage of choices in which the subjects followed the link information provided (by choosing the link that had less level of congestion) versus the choices that did not follow the information (by choosing the link that had a higher level of congestion). The percentages presented in Figure 4 are categorized by the information type. The highest difference occurred in scenario 4, where en-route information is provided. Scenario 5, where an advised route from the driver's position to the destination is also provided, holds the second highest difference. The first three scenarios aggregated together, where no en-route information is provided, have the least difference. Using the χ^2 test, the number of choices that followed the information provided occurred in scenarios #4 and #5 were found significantly greater than those of the other 3 scenarios ($\chi^2 = 10.55$, p-value < 0.001). This underlines the benefits of en-route information on short-term driver's route choices (link choices).

7.1. *Modeling short-term en-route link travel time*

The 3086 link choices were the data used in this model. The mixed linear model with repeated observations was employed. At every link choice, the delay on the link that has not been chosen minus that of the chosen link

Table 2. The estimated parameters of the unstructured correlation matrix.

Choice #	1	2	3	4	5	6	7	8	9	10
1	1	0.388	0.031	0.242	0.060	0.332	0.073	0.140	0.096	0.061
2		1	0.020	0.030	0.220	0.237	0.091	0.018	0.083	0.130
3			1	0.273	0.129	0.175	0.078	0.134	0.025	0.315
4				1	0.138	0.117	0.114	0.107	0.045	0.082
5					1	0.078	0.005	0.073	0.001	0.189
6						1	0.437	0.060	0.293	0.048
7							1	0.041	0.151	0.008
8								1	0.169	0.100
9									1	0.037
10										1

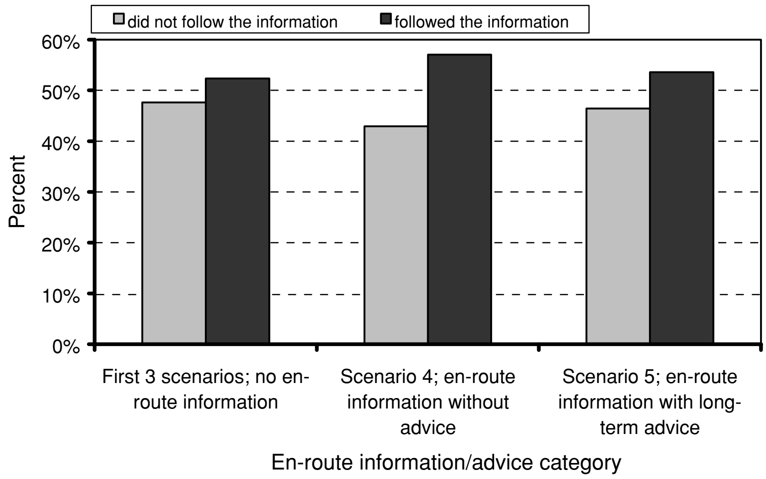

Figure 4. Percent of link choices that followed or did not follow versus information/advice category.

represents the dependent variable of the model. This represents the amount of the travel time saving if the subject chooses the link with less delay. It also represents the amount of extra delay if the subject chooses the link with more delay. In both cases, the response variable is the difference in delay between two links not the absolute delay of the chosen link. The delay of a link, at a certain movement, was taken as equal to its actual travel time minus its free flow travel time. The value of the response variable represents the travel time saving that a subject achieved by following the short-term traffic information provided. A negative response value means that the subject was delayed (for the short-term) by choosing the more congested link. While if he/she had chosen the other link then the delay would have been reduced. The explanatory variables are:

1. Information familiarity; 1 if subject uses pre-trip and/or en-route traffic information usually or everyday, "0" otherwise
2. Information provision; 1 for scenario 4 where en-route information is provided without advice, 0 otherwise
3. Same color; 1 if the two coming links had the same color (qualitative congestion level), 0 otherwise. This variable tests the effect of qualitative versus quantitative information
4. System learning; 1 for the second 5 trial days of the simulator, "0" for the first 5 trial days
5. Number of movements from the origin

Table 3 shows the parameter estimates with a comparison between the CS and the independent case. The maximum number of repeated choices per subject was 71. The number of unknown and necessarily estimated parameters in the covariance matrices were [1], [2], [71 (72)/2 = 2,556] for the Independent, CS, and UN structure, respectively. Therefore, the UN correlation matrix was not appropriate for this model. Shown in Table 3, the likelihood ratio test (between the convergent versus restricted models) proved that the explanatory variables could statistically explain the response variable. The independent case is a special case of the CS correlation structure (Littell et al. 1996), then a restricted likelihood ratio test with test statistic equal to twice the difference between the two log likelihoods follows a Chi-squared distribution with degrees of freedom equal to the difference in the number of parameters to be estimated in the covariance matrices. The χ^2 test results showed that the CS structure was the best. AIC agreed with these results as the CS structure has the smallest AIC value.

The modeling results showed that the en-route short-term information (link) has a good chance to be used and followed. Drivers familiar with the traffic information had relatively less short-term delay or more short-term travel time saving under ATIS. Providing en-route short-term traffic infor-

Table 3. Results of short-term travel time model.

	Independent		CS	
	Coeff.	t-static	Coeff.	t-static
Intercept	1.969	3.47	1.893	3.56
Info-familiarity; 1 if subject uses pre-trip and/or en-route traffic information usually or everyday, 0 otherwise	0.899	3.53	0.958	3.38
Info-provision: 1 for scenario 4 where en-route information is provided without advice, 0 otherwise	0.560	2.67	0.590	2.78
Same color: 1 if the two coming links were with the same color (qualitative congestion level), 0 otherwise	−0.661	−4.11	−0.659	−4.29
System learning: 1 for the second 5 trial days of the experiment, 0 for the first 5 trial days	2.967	60.08	2.962	61.17
Number of movements since the origin	4.361	43.45	4.319	43.89
Interaction term				
Number of movements since the origin * System-learning	4.834	42.36	4.783	42.82
Summary statistics				
Sample size = 3086				
Log likelihood at zero [L (c)]	−10478.4		−10477.9	
Log likelihood at convergence [L (β)]	−8978.5		−8972.1	
LR statistic = 2 [L(β)−L (c)], D.O.F = 6	2999.8		3011.5	
AIC	17957.3		17948.1	

Chi-square test results to compare between the three covariance structures used:
LR (CS versus Indep.); $\chi^2_{\text{stat.}} = 12.8$; D.O.F = 2−1 = 1; Favors CS.

mation was shown to reduce travel time. This indicates the significance of en-route information in short-term choices. Advice was not important since it addresses the long-term (whole route). Learning and being familiar with the system that provides the information also reduces the short-term delay. When being away from the origin, i.e. close to the destination (presented by the number of movements since the origin), drivers had less delay. This means that drivers are more likely to follow the short-term information when they get closer to the destination. It is worth mentioning that system learning and the number of movements have high t-statistics indicating small standard errors given the relatively large coefficient estimates. These two variables are highly significant in the presence of the other factors. When the two coming links had the same qualitative level of congestion drivers experienced higher delays. It can be concluded that the qualitative information is more likely to be used than the quantitative information. Other factors including; age, gender, income, driving experience, familiarity with the network, weather conditions, and frequency use of expressways are tested and found to be uncorrelated with the response variable. Unlike most of the existing studies including the first model of this paper, this model analyzes en-route link travel time under different types and levels of ATIS.

8. Conclusions

This paper investigates the factors that affect driver's accessing and benefiting from real-time pre-trip and en-route traffic information with/without advice. Origin-to-destination and node-to-node travel time were modeled separately. A travel simulator with realistic network and real historical volumes was used to collect dynamic route choice data. The simulator provided five different types and levels of ATIS. The Mixed Linear Model with the repeated observation's technique was used for both models. The correlation was found to be significant in both models, which underlines the importance of accounting for correlation in similar studies. Different variance-covariance structures (Independent, CS, and the UN) were used and compared to ensure the validity of the statistical analysis.

The results showed that as the level of information increases (adding en-route to the pre-trip and advice to the advice-free-information) the average travel time decreases. Drivers saved up to 44.7% of the overall travel time when equipped with pre-trip and en-route information with advice, relative to no information. The results showed that driver's who are familiar with traffic information, the network, and the system that provides the information had relatively less delay. The link choice analysis showed that providing en-route short-term traffic information reduces the en-route delay. Learning and being familiar with the system that provides the information and being close to the destination increase the usage of the en-route information and therefore reduce the en-route delay. The qualitative information is more likely to be used than the quantitative information.

This paper underlines the importance of modeling the temporal correlation between repeated choices made by the same traveler. Considering different correlation structures is also important and depends on the nature of the existing correlation and on the sample size. The paper addresses the benefits and use of ATIS in route choice in a microscopic level (one driver at a time). Future work should address the issue of market penetration which will likely affect the performance of the network.

References

Abdalla M (2003) *Modeling Multiple Route Choice Paradigms Under Different Types and Levels of ATIS Using Correlated Data, Ph.D. dissertation, University of Central Florida.*

Abdel-Aty M & Abdalla F (2002) *Design and development of a computer simulation tool to collect data that support hybrid travel models under ATIS.* Montreal, Quebec, Canada: Proceedings of the Annual Conference of the Canadian Society for Civil Engineering.

Abdel-Aty M, Kitamura R & Jovanis P (1997). Using stated preference data for studying the effect of advanced traffic information on driver's route choice. *Transportation Research Part C* 5:(1), 39–50.

Abdel-Aty M, Kitamura R & Jovanis P (1995). Investigating effect of travel time variability on route choice using repeated measurement stated preference data. *Transportation Research Record* 1493: 39–45.

Abdel-Aty M, Vaughn K, Jovanis P, Kitamura R, & Mannering F (1994a) Impact of traffic information on commuters' behavior: empirical results from Southern California and their implications for ATIS. *The Proceedings of the 1994 Annual Meeting of IVHS America*, pp. 823–830.

Abdel-Aty M, Kitamura R, Jovanis P, & Vaughn K (1994b) *Investigation of Criteria Influencing Route Choice: Initial Analysis Using Revealed and Stated Preference Data.* Report No: UCD-ITS-RR-94–12, Davis: California University.

Arnott R, De-Palma A & Lindsey R (1990). Departure time and route choice for the morning commute. *Transportation Research Part B* 24: 209–228.

Bonsall P & Parry T (1991). Using an interactive route-choice simulator to investigate driver's compliance with route guidance advice. *Transportation Research Record* 1306: 59–68.

Chen P, Srinivasan K & Mahmassani H (1999). Effect of information quality on compliance behavior of commuters under real-time traffic information. *Transportation Research Record* 1676: 53–60.

Chen W. & Jovanis P (2003) *Analysis of Driver En-Route Guidance Compliance and Driver with ATIS using a Travel Simulation Experiment.* Washington, DC: presented in the 81st Transportation Research Meeting.

Delvert K (1997) *Heterogeneous Agents Facing Route Choice: Experienced versus Inexperienced Tripmakers.* Presented at IATBR '97, Austin, TX.

Garrido R & Mahmassani H (2000). Forecasting freight transportation demand with the space-time multinomial probit model. *Transportation Research Part B* 34:(5), 403–418.

Gopinath D (1995) *Modeling Heterogeneity in Discrete Choice Processes: Application to Travel Demand* Ph.D. Thesis, MIT.

Guerin L & Stroup W (2000) *A simulation study of evaluation PROC MIXED analysis of repeated measures data*. Manhattan, KS: Kansas State University: Proceedings of the 12th Annual Conference on Applied Statistics in Agriculture.

Jou R & Mahmassani H (1998) *Day-To-Day Dynamics of Urban Commuter Departure Time and Route Switching Decisions: Joint Model Estimation*. New York: Elsevier: Travel Behavior research.

Khattak A, Schofer J, & Koppelman F (1995) Effect of traffic information on commuters' propensity to change route and departure time. *Journal of Advanced Transportation* 29, 193–212.

Kitamura R & Bunch D (1990) *Heterogeneity and State Dependence in Household Car Ownership: A panel Analysis Using Ordered-Response Probit Models with Error Components Transportation and Traffic Theory*. Oxford: Elsevier 477–496.

Koutsopoulos K, Duse-Anthony Y, Fisher D & Duffy S (2000). The framing of drivers' route choice when travel time information provided under varying degrees of cognitive load. *Human Factors* 42:(3), 470–481.

Littell R, Stroup W, & Freund R (1996a) *SAS for Linear Models, Fourth edition. ISBN 0-387-95027-3, Cary, NC: SAS Institute Inc., USA.*

Liu Y & Mahmassani H (1998). Dynamic aspects of departure time and route decision behavior under advanced traveler information systems (ATIS): modeling framework and experimental results. *Transportation Research Record* 1645: 111–119.

Louviere J & Woodworth G (1983). Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research* XX: 350–367.

Mahmassani H & Liu Y (1999). Dynamics of commuting decision behavior under advanced traveler information systems. *Transportation Research Part C* 7:(2), 91–107.

Mahmassani H & Hu T (1997). Day-to-day evolution of network flows under real-time information and reactive signal control. *Transportation Research Part C* 5: 51–69.

Mannering F, Kim S, Barfield W & Ng L (1994). Statistical analysis of commuters' route, mode, and departure time flexibility. *Transportation Record C* 2:(1), 35–47.

Mannering F (1987) Analysis of the impact of interest rates on automobile demand. *Transportation Research Record* 1116: 10–14.

Morikawa T (1994) Correcting state dependence and serial correlation in the RP/SP combined estimation method. *Transportation Journal* 21(2).

Pallottino S & Grazia M (1988) Shortest Path Algorithms in Transportation Models: Classical and Innovative Aspects. In Marcotte P and Nguyen S (eds) *Equilibrium and Advanced Transportation Modeling* (pp. 245–28).

Reiss R, Gartner N & Cohen S. (1991). Dynamic control and traffic performance in a freeway corridor: a simulation study. *Transportation Research Part A*, 25:(5), 267–276.

SAS/STAT (2003), Technical Support Resources, Version 8.2.

Searle S, Casella G & McCulloch C (1992) *Variance Components*.. New York: John Wiley & Sons Inc.

Searle S (1971) *Linear Models* New York: John Wiley & Sons Inc..

Sengupta R, & Hongola B (1998) Estimating ATIS benefits for the Smart Corridor Partners for Advanced Transit and Highways, Research report, UCB-ITS-PRR-98-30, In: *Institute of Transportation Studies*, Berkeley: University of California..

Stokes M, Davis C & Koch G (2000) *Categorical data analysis using the SAS system, second edition,* Cary, NC: SAS Institute.

Vaughn K, Kitamura R & Jovanis P (1995a). Experimental analysis and modeling of advice compliance: results from advanced traveler information system simulation experiments. *Transportation Research Record* 1485: 18–26.

Vaughn K, Reddy P, Abdel-Aty M, Kitamura R & Jovanis P (1995b). Route choice and information use: initial results simulation experiment. *Transportation Research Record* 1516: 61–91.

Verbeke G, & Molenberghs G (2000) Linear Mixed Models for Longitudinal Data. In: *Springer series in Statistics*, ISBN 0-387-95027-3, Springer-Verlag New York, Inc.

Wolfinger R & Chang M (1996) *Comparing the SAS®* GLM and MIXED Procedures for repeated observations. Cary, NC: SAS institute INC.

Wunderlich K (1996) An assessment of pre-trip and en route ATIS benefits in a simulated regional urban network. Proceedings of the third world congress on Intelligent Transport Systems at Orlando.

Yamamoto T, Fujii S, Kitamura R & Yoshida H (2000). An analysis of time allocation, departure time and route choice behavior under congestion pricing. *Transportation Research Record* 1725: 95–101.

Zhao S, Harata N & Ohta K (1996) *Assessing driver benefits from information provision: a logit model incorporating perception band of information*. London, England: Proceedings of the 24th annual Passenger Transport Research Conference.

About the authors

Mohamed Abdel-Aty is an Associate Professor of Civil Engineering at the University of Central Florida (UCF). His main research interest is in the areas of traffic safety, travel demand analysis, and ITS. He has more than 100 published papers. He is a member of two TRB committees, and co-chair of the subcommittee on ATIS. He is a member of the Editorial Advisory Board of Accident Analysis and Prevention and the editorial board of the ITS Journal. He is the 2003 UCF Distinguished Researcher.

M. Fathy Abdalla completed his Ph.D. in Transportation Engineering from the University of Central Florida in 2003. He is a lecturer in transportation at Zagazig University in Egypt. His main research interests are in the areas of travel demand.

COPYRIGHT INFORMATION

TITLE: Modeling travel time under ATIS using mixed linear models
SOURCE: Transportationcavhlth Jron 33 no12102/12 Ja 20060052004

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited. To contact the publisher: http://www.springerlink.com/content/1572-9435/