

# UC Berkeley

## Earlier Faculty Research

### Title

A Structural Model with Discrete-Choice Variables for Predicting Enroute Behavior under ATIS

### Permalink

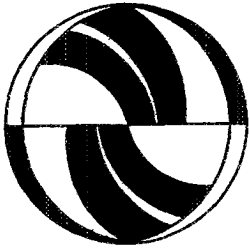
<https://escholarship.org/uc/item/99k5h6jq>

### Authors

Adler, Jeffrey L.  
Golob, Thomas F.  
McNally, Michael G.

### Publication Date

1993-08-01



**A Structural Model with Discrete-Choice  
Variables for Predicting Enroute Behavior  
under ATIS**

Jeffrey L. Adler  
Thomas F. Golob  
Michael G. McNally

Working Paper  
UCTC No. 174

**The University of California  
Transportation Center**

University of California  
Berkeley, CA 94720

**The University of California  
Transportation Center**

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, Irvine, and Los Angeles, the Institute of Urban and Regional Development at Berkeley, and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses. Faculty and students on other University of California campuses may participate in

Center activities. Researchers at other universities within the region also have opportunities to collaborate with UC faculty on selected studies.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there.

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. It also publishes *Access*, a magazine presenting summaries of selected studies. For a list of publications in print, write to the address below.

**DISCLAIMER**

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.



**University of California  
Transportation Center**

108 Naval Architecture Building  
Berkeley, California 94720  
Tel: 510/643-7378  
FAX: 510/643-5456

The contents of this report reflect the views of the author who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the U.S. Department of Transportation. This report does not constitute a standard, specification, or regulation.

**A Structural Model with Discrete-Choice Variables for  
Predicting Enroute Behavior under ATIS**

Jeffrey L. Adler

Department of Civil and Environmental Engineering  
Rensselaer Polytechnic Institute  
Troy, NY 12180

Thomas F. Golob

Institute of Transportation Studies  
University of California at Irvine  
Irvine, CA 92717

Michael G. McNally

Institute of Transportation Studies and  
Department of Civil Engineering  
University of California at Irvine  
Irvine, CA 92717

*Working Paper  
August 1993*

UCTC No 174

The University of California Transportation Center  
University of California at Berkeley



# **A STRUCTURAL MODEL WITH DISCRETE-CHOICE VARIABLES FOR PREDICTING ENROUTE BEHAVIOR UNDER ATIS**

**Jeffrey L. Adler**

Department of Civil and Environmental Engineering  
Rensselaer Polytechnic Institute  
Troy, New York 12180-3590

**Thomas F. Golob**

Institute of Transportation Studies  
University of California Irvine, Irvine, CA 92717-3600

**Michael G. McNally**

Institute of Transportation Studies  
and Department of Civil Engineering  
University of California Irvine, Irvine, CA 92717-3600

## **ABSTRACT**

Increasing efforts in Advanced Traveler Information Systems has emphasized the need to develop more robust models of enroute driver behavior. The complexity in modeling driver behavior stems from the need to capture the day-to-day dynamics of choice, model diversion and active information acquisition, and account for individual preferences and needs. Previous papers by the authors discussed a conflict arousal methodology for modeling driver behavior in the presence of real-time information. It was posed that observed changes in enroute driver behavior, characterized by diversion or active information acquisition, are evoked through latent factors of conflict arousal and motivation. In-laboratory interactive simulation was used to collect enroute behavior data. This paper presents a structural equation approach for modeling enroute driver behavior and explaining causal factors of diversion and information acquisition as well as the interrelationship between these observed behaviors.



## INTRODUCTION

Several key factors are known to influence enroute choice including, but not limited to, perception and knowledge of current path conditions (delay, expected travel time, congestion level), knowledge (or perception) of the existence of alternate routes and travel conditions on these routes, affinity toward taking risks, and certainty of meeting travel goals or expectations. In addition to these variations across a population, driving behavior also varies for individuals based on trip characteristics such as trip type, time of day, and other considerations.

The dynamics of enroute driver behavior have been modeled as an iterative process through which drivers are constantly assessing the current state of the travel process and adjusting their behavior in response to changes in the system (Ben-Akiva et al, 1991; Adler et al., 1992a; and Khattak et al, 1993). Based on this iterative formulation researchers have been trying to estimate under what conditions drivers consider adapting enroute behavior by diverting from the current path or by acquiring additional information to assist in assessing network conditions.

A common theory on diversion behavior (Khattak et al, 1993, Adler et al., 1992) holds that drivers react to increased frustration and anger at deteriorating travel conditions. Thresholds, measuring drivers' tolerance to increased frustration, can be used to predict diversion behavior. Increased difference between perceived and expected travel conditions evoke stronger feelings of frustration among drivers which in turn impel drivers to changes their behavior in hopes of reducing these anxieties. Thresholds of expectations may vary across individuals or for the same driver depending on temporal and spatial considerations.

In previous papers (Adler et al., 1992), a more general approach to thresholds and enroute behavior was developed. This conflict theoretical approach, based on fundamental principles of psychology and human choice behavior, characterized enroute behavior as a three stage process of arousal, motivation, and response.

The conflict theoretical approach is based on three assertions of behavior:



- (i) Drivers are goal attainers; all actions undertaken as part of the travel process are designed to help drivers better their chances of meeting their travel goals.
- (ii) Changes in behavior are triggered by an increased dissonance between drivers' current perceived ability to achieve travel goals versus their desired goal attainment state grows
- (iii) Changes in behavior are effectuated through drivers' perceived ability to significantly improve the chances of goal attainment or reducing the anxiety and frustration caused by the increased dissonance of goal attainment.

It is assumed that drivers specify a set of travel objectives during the pre-trip planning phase. This set of objectives reveals a driver's perception of travel conditions and, under expected travel conditions, will optimize the driver's utility of travel. Based on the expected travel conditions, an initial route will be selected that will optimize the utility defined by the set of objectives. Enroute, as long as travel conditions remain normative, the initial route choice will satisfy the objective space and optimize driver's utility of travel. Changes in enroute behavior should only be expected under conditions of increased congestion and uncertainty of perception.

Expounding on the concept of expectation thresholds, it is theorized that drivers become aroused to adapt enroute behavior as their conflict level increases and exceeds a tolerable threshold. Conflict is modeled as a latent factor measuring increased anxiety and frustration experienced as drivers perceive an increased inability to attain their travel objectives. The difference between perceived and desired travel objectives is used to measure driver dissatisfaction with the current travel path. When conflict exceeds a tolerable threshold, changes in enroute behavior are not guaranteed. It is necessary that drivers recognize potential alternative paths. Moreover, drivers must have significant motivation to adapt their behavior (e.g., drivers will divert only when an acceptable marginal increase in utility is to be gained). Appropriate responses to increased conflict arousal and motivation can involve one or more of the following actions: diversion, information acquisition, reevaluation of travel objectives, or do nothing.

The conflict arousal and motivation approach differs from many traditional approaches to modeling dynamic driver behavior in that latent, or unobserved factors, are directly incorporated in the theory. Arousal and motivation cannot be directly inferred from single observable input or output variables but rather are situation factors that may involve several factors working together to influence enroute driver behavior. As a result, structural equations models provide a robust approach to modeling latent factors and for directly testing various hypotheses of relationships between observed inputs and outputs.

A previous study of diversion behavior, using in-laboratory interactive simulation for data collection (Adler et al., 1993), was performed using binomial logit analysis. Although binary logit is generally unsuited for modeling latent factors, the analyses were undertaken to identify some inherent relationships between travel attributes and diversion behavior. It was found that primary and secondary diversions are influenced by driver's familiarity with congestion levels and layout of current and alternate paths, availability of ATIS, distance from the destination, and perceived travel speeds. It was also concluded that after drivers initially change from a primary route to a secondary route, there is a greater willingness to change routes again - almost a real-time wayfinding behavior. The results of this preliminary analysis provided the basis for using structural equation techniques for testing the conflict arousal and motivation theory.

This paper describes a series of models that were developed with structural equation approaches to model enroute driver behavior and test the theories of conflict arousal and motivation described above. The next section provides a brief overview of structural equation modeling. This is followed by a description of the case study and data collection process. The latter parts of the paper describe how structural equations were used for analysis, the results of the estimation process, and provide some insights for further research.

## **THE CASE STUDY**

Data for this research was collected during a case study performed at the University of California Irvine using in-laboratory interactive microcomputer simulation. FASTCARS

(Freeway and Arterial Street Traffic Conflict Arousal and Resolution Simulator), a simulator that emulates the driving process and incorporates three types of advanced traveler information systems technologies (variable message sign, highway advisory radio, and in-vehicle navigation system), was developed for data collection purposes. A brief overview of the experiment is documented below; for a full description of FASTCARS and the case study, please refer to Adler et al., (1992a, 1992b, 1993)

A special event trip was simulated with FASTCARS. Participants were invited to play the simulator and take a trip through a hypothetical network from a present origin to a present destination. FASTCARS leads players through the pre-trip planning, enroute travel, and post-trip evaluation phases. During the pre-trip phase, player are shown pictures of the network and asked to plan their initial route. Players performances are evaluated based on a set of five travel goals. A logistic function is used to normalize each goal to a scale of 0-100 points. To identify an individual's travel objectives, players are asked to distribute 100 weighting points to the goal set. At the conclusion of the program a player's score is determined from a linear additive function in which the values attained for the five goals are multiplied by the assigned weights and summed together to provide a maximum possible score of 10000 points.

To model various levels of player familiarity, three categories of 'network experience' were created: novice, intermediate, and expert. Separate sets of maps with varying detail were created for each level of experience. Novice players were provided maps containing only schematics and distanced for a few major freeways in the network. Intermediate maps provided information about the entire freeway system (layout, distances, and expected speeds) as well as major arterials (layout and distances). Maps for expert players not only contained detailed schematics of the entire network but also provided information concerning expected speeds and probability of incidents. Initially, participant are randomly assigned novice or intermediate status. After playing a number of trials, players may be graduated to higher levels of experience

Enroute travel is modeled in real-time and players attempt to navigate their vehicle through the network. Active choices that can be made by players include lane changing, road

changing, acquiring information from the highway advisory radio or in-vehicle navigator, and changing the weighting scheme on the goal set. Maps of the network may also be viewed at any time during travel. At each simulation event, the player's actions and the state of the network are recorded. Throughout the trip players may experience various levels of traffic congestion and may passively acquire traffic or network-related information conveyed through variable message or other road-side signs.

Players, recruited from a university population, participated in at least two and up to a maximum of ten trials (each trial was one trip); a total of 108 trials were collected. The initial session was conducted under guidance of an instructor who explained the workings of the simulator; successive trials were performed at players' convenience.

## **DATA SET AND RESEARCH HYPOTHESES**

For the structural equation analysis, each trip was divided into four sections based on travel time. The total travel time of the trip was divided into four sections, each assigned a quarter of the travel time. To maintain the model's integrity it was important to have the same number of observations for each player. It was determined that dividing trips into four sections produced the best cross section of data. Each section averaged about ten minutes of travel time and there was an even spread of output events. Data sets with greater than four sections produced too many sections with few observed events; data sets with less than four sections produced too few events with limited distribution of observed outputs. For the 108 trials collected, a total of 432 new records were classified for the structural equation analysis.

For each trip section, fourteen variables were calculated. These variables, listed in Table 1, define the set of endogenous and exogenous variables considered for the analysis. The three observed outputs (**Divert**, **HAR**, **IVNS**) defined the possible player actions in each trip section. These variables were assigned [0,1] values with 1 indicating that the output was observed during the specified trip section. The remaining eleven variables were designated observed inputs. In a series of category models performed on the data set (see Adler et al., 1993), these variables were



seen as significant factors of travel behavior. Average values were calculated for these inputs except for the variables **enddist** and **VMS**.

**Enddist** measures the distance from the destination. It is hypothesized that diversion behavior is influenced by a driver's proximity to the destination. Since drivers experienced trips with varied distances, this variable was assigned values [1-4] indicating the time slice of the variable. Lower values represent further distance from the destination. **VMS**, indicating the presence of variable message signs observed during the time slice, was assigned the maximum message level. Four categories of messages were emulated during the case study ranging in severity from [1 (no congestion) - 4 (major incident)]. Therefore, **VMS** was coded [0-4] -- 1-4 indicating message sign observed and 0 if the driver did not encounter a message sign during that time slice.

Driver's familiarity with network layout and travel conditions, both on selected and alternate paths, has a significant influence on behavior. Therefore, level of familiarity with network configuration and conditions is one of the variables controlled for in the case study. Participants in the case study were randomly assigned a player level (1-novice, 2-intermediate, 3-expert) and maps, drawn with varied levels of network configuration and traffic condition information were provided for guidance. Novice players were provided with maps having the least detail and maps given to expert players had the most detail. In effect, maps and player levels were used in the case study as proxies for experience.

The ratio of player level to map level of the link (**mtl**, **matl**) was calculated as an indicator of network experience. Links in the network were assigned values depending on the lowest level maps on which they appeared. For example, links visible to all players were coded 1 and links visible only to experts were coded 3. Values greater than unity indicate links on which a player's experience is greater than or equal to the lowest level of map on which the link is displayed.

It is hypothesized that the conflict arousal and motivation latent processes for diversion and information acquisition are influenced by groups of variables. Initiation of the conflict arousal stage involves a driver's perception of changing travel conditions and ability to predict arrival time

to the destination. Variables that can capture this perception include perceived speed, variable message signs, and distance from destination. Motivation to change behavior is impacted by other variables that reflect link or path specific attributes such as road types (e.g., arterial vs. freeway) and familiarity with layout and travel conditions.

Two sets of analyses were performed using the LISREL model. The first set illustrates how structural equations can be used for multivariate regression analysis. The second approach incorporates latent variables in a model assuming error in the y-variables only. A set of  $\eta$  latent endogenous variables are estimated between the x and y variables.

### **STRUCTURAL MODEL WITH DISCRETE-CHOICE VARIABLES**

The theoretical concepts of interest are captured directly by three endogenous discrete-choice variables: route diversion (DIVERT), propensity to use highway advisory radio (HAR), and propensity of using in-vehicle navigation systems (IVNS).

It is assumed that each of these endogenous dichotomous variables is the observed consequence of a probability process that potentially depends on all of the other exogenous and endogenous variables in the system. For each dichotomous variable  $y_i$ , ( $i = 1,2,3$ ), there is a latent discrete-choice variable  $y_i^*$ , ( $i = 1,2,3$ ) that is normally distributed with mean zero and variance one. The relationship between each observed and latent discrete-choice endogenous variable is:

$$\begin{aligned} y_i &= 1 \text{ if } y_i^* \geq \alpha_i \\ y_i &= 0 \text{ otherwise} \end{aligned} \tag{1}$$

where  $\alpha_i$  is the threshold value of the cumulative normal distribution function corresponding to the marginal distributions of the population over the two categories.

These thresholds can be estimated using the probit model (Goldberger, 1964; Maddala, 1983). The probability of observing  $y_i = 1$ , conditional on the exogenous variables in the system is given by:

$$P(y_i = 1 | x) = P(y_i^* \geq \alpha_i) = 1 - \Phi(\alpha_i - \omega_i' x) \quad (2)$$

where  $\Phi$  denotes the standard cumulative normal distribution and  $\omega_i$  is a vector of reduced-form regression coefficients defining the conditional mean of  $\alpha_i$  given  $x$ . The likelihood function for this binomial process corresponding to the above probability is:

$$L = \prod_{j=1}^n [\Phi(\alpha_i - \omega_i' x_j)]^{y_{ij}} [1 - \Phi(\alpha_i - \omega_i' x_j)]^{(1-y_{ij})} \quad (3)$$

where  $y_{ij}$  is the value of  $y_i$  observed for individual  $j$  and  $x_j$  is the vector of exogenous variables for individual  $j$ . The log-likelihood is given as:

$$\log L = \sum_{j=1}^n y_{ij} \log \Phi(\alpha_i - \omega_i' x_j) + \sum_{j=1}^n (1 - y_{ij}) \log [1 - \Phi(\alpha_i - \omega_i' x_j)] \quad (4)$$

The  $\alpha_i$  and  $\omega_i$  values that maximize the log-likelihood function provide the distribution of  $y_i^*$  that is least surprising.

The structural equation model in terms of the  $y_i^*$  discrete-choice variables is:

$$y_i^* = B y_i^* + \Gamma x + \zeta \quad (5)$$



The hypothesis to be tested is specified in terms of the B and  $\Gamma$  structural parameter matrices in equation (5). The B matrix, capturing the direct effects between the endogenous discrete-choice probabilities specified to be:

$$B = \begin{bmatrix} 0 & \beta_{12} & 0 \\ 0 & 0 & 0 \\ 0 & \beta_{32} & 0 \end{bmatrix} \quad (6)$$

which implies that  $y_2^*$ , the probability of seeking highway advisory radio, effect both  $y_1^*$ , the probability of diversion, and  $y_3^*$ , the probability of seeking in-vehicle navigation information. All other possible direct effects between these three discrete-choice probabilities are specified to be negligible: the probability of diversion has no influence on either the probabilities of seeking HAR or IVNS; and the probability of seeking IVNS has no influence on either diversion or seeking HAR.

The  $\Gamma$  matrix, capturing direct effects on the eleven exogenous variables on the three endogenous discrete-choice probabilities is specified as:

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 & \gamma_{15} & 0 & 0 & 0 & \gamma_{19} & 0 & \gamma_{111} \\ \gamma_{21} & \gamma_{22} & 0 & \gamma_{24} & 0 & \gamma_{26} & 0 & \gamma_{28} & 0 & \gamma_{210} & 0 \\ \gamma_{31} & 0 & 0 & \gamma_{34} & 0 & 0 & \gamma_{37} & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7)$$

This specification implies that there are direct effects between certain exogenous variables and endogenous discrete-choice probabilities.

Finally,  $\zeta$  is a 3 x 1 vector of errors (unique, or unexplained portions) of the  $y^*$  variables.

Estimation of a structural equation model with discrete-choice endogenous variables is accomplished by applying the generalized least squares moment-estimation method to a correlation matrix formed by tetrachoric correlations between the endogenous variables, and



polyserial correlations between the endogenous variables and the exogenous variables. Tetrachoric correlation coefficients are estimates of the unobserved correlations between two unobserved probit variables, say  $y_1^*$  and  $y_2^*$ , responsible for the two dichotomous variables,  $y_1$  and  $y_2$  according to equation (1). A cross-tabulation of the two observed dichotomous variables produces cell frequencies  $N_{ij}$ ,  $i = 1, 2$  and  $j = 1, 2$ . The log-likelihood resulting from these observed cross-tabulation (contingency table) frequencies is then

$$\log L = \text{constant} + \sum_{i=1}^c \sum_{j=1}^d N_{ij} \log(\pi_{ij}) \quad (8)$$

where

$$\pi_{ij} = \Phi_2(\alpha_1, \alpha_2) - \Phi_2(\alpha_1, -\infty) - \Phi_2(-\infty, \alpha_2) \quad (9)$$

and  $\Phi_2$  denotes the bivariate normal distribution function with correlation  $\rho$ . The best estimate of the tetrachoric correlation is the  $\rho$  value that maximizes this log-likelihood of observing the cross-tabulation frequencies, given the two thresholds  $a_1$  and  $a_2$  found by maximizing equation (4) in the first step of the estimation method (Kirk, 1973). Polyserial correlation coefficients are similarly computed by as the unobserved correlation between a probit latent variable and a continuous variable (Olsson, et al., 1973).

The free structural parameters in the B and  $\Gamma$  matrices, denoted by the vector  $\hat{\theta}$ , are estimated by making the model-implied covariance matrix,  $\Sigma(\hat{\theta})$  as close as possible to the sample covariance matrix, S, where S is composed of tetrachoric, and polyserial correlation coefficients, depending upon variable type. It is not appropriate to use the popular *FML* maximum-likelihood fitting function, because the assumptions underlying this method do not hold for dichotomous endogenous variables. Maximum likelihood estimation in this case will yield consistent estimates but incorrect standard errors (t-statistics) and  $X^2$  statistics.

The best estimation method is weighted least squares (WLS). The fitting function for WLS is

$$F_{WLS} = [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})]' \mathbf{W}^{-1} [\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})] \quad (10)$$

where  $\mathbf{s}$  is a  $\left[\frac{1}{2}(p+q)(p+q+1)\right]$  vector of tetrachoric, and polyserial correlation coefficients for all pairs of ( $p$ ) endogenous and ( $q$ ) exogenous variables,  $\boldsymbol{\sigma}(\boldsymbol{\theta})$  is a vector of model-implicated correlations for the same variable pairs, and  $\mathbf{W}$  is a  $\left[\frac{1}{2}(p+q)(p+q+1)\right] \times \left[\frac{1}{2}(p+q)(p+q+1)\right]$  positive-definite weight matrix. Minimizing  $F_{WLS}$  implies that the parameter estimates are those that minimize the weighted sum of squared deviations of  $\mathbf{s}$  from  $\boldsymbol{\sigma}(\boldsymbol{\theta})$ . This is analogous to weighted least squares regression, but here the observed and predicted values are correlations rather than raw observations.

The best choice of the weight matrix is a consistent estimator of the asymptotic covariance matrix of  $\mathbf{s}$ :

$$\mathbf{W} = ACOV(s_{ij}, s_{gh}) \quad (11)$$

Under very general conditions

$$\mathbf{W} = \frac{1}{N} (\sigma_{ijgh} - \sigma_{ij} \sigma_{gh}) \quad (12)$$

is a consistent estimator, where  $\sigma_{ijgh}$  denotes the fourth-order moments of the variables around their means, and  $\sigma_{ij}$  and  $\sigma_{gh}$  denote covariances. Brown (1974, 1984) demonstrated that  $F_{WLS}$  with such a weight matrix will yield consistent estimates  $\hat{\boldsymbol{\theta}}$  which are asymptotically efficient with correct  $ACOV(\hat{\boldsymbol{\theta}})$  (leading to correct parameter t-statistics) and correct  $X^2$  test values. These properties hold for very general conditions, and consequently such  $F_{WLS}$  estimators are known as arbitrary distribution function, or asymptotically distribution free (ADF) estimators.



For the hypothesized model, the estimation results are shown in Figures 1 and 2 and Table 3. This model is significant at  $p=0.876$  with a chi-square of 12.223 and 19 degrees of freedom. It also has an Adjusted Goodness of Fit Index of 0.990 and Root Mean Square Residual of 0.0303.

Other possible alternative hypotheses involving parameters of the Beta matrix ( $\beta_{21}$ ,  $\beta_{31}$ ,  $\beta_{13}$ ,  $\beta_{23}$ ) were tested and shown that they could be rejected. Table 2 summarizes these results. The first model listed in the table has significant parameters but does not fit as well, it has a high chi-square and large RMSR. The other three models presented in the table have poor fits. In each case, there are betas that have non-significant t-values and, as indicated by LISREL's modification indices tables, there were potentially better variable combinations.

### **STRUCTURAL EQUATION MODEL WITH LATENT DISCRETE-CHOICE FACTORS**

A powerful feature of structural equation models is that they can be used to estimate latent factors. The conflict arousal and resolution concepts advanced in this research hypothesize that there are latent factors that exist between driver perception and action. These steps of arousal and motivation dictate which actions will follow.

Based on the theoretical model of conflict arousal and motivation, a second set of models was estimated using structural equations with measurement models. The full equation including the latent discrete-choice factors is given by two equations. The first denotes the full model; the second implies the measurement model for the latent factors on the endogenous variables:

$$\begin{aligned} \eta &= B\eta + \Gamma x + \zeta \\ y^* &= \Lambda\eta + \varepsilon \end{aligned} \tag{13}$$

The  $\Lambda$  matrix contains the coefficients between the latent factors and the endogenous variables. The  $B$  matrix is now used to capture the direct effects between the latent discrete-

choice factors, and the  $\Gamma$  matrix captures direct effects of the eleven exogenous variables on the latent discrete-choice factors.

The hypotheses to be tested in this model suggest that latent concepts of arousal and motivation can be represented by four latent discrete-choice factors. Each of the first three factors lead directly into a single endogenous discrete-choice variable; these represent the motivation factor. The arousal phenomenon is represented by a single latent factor that is directly effected by a subset of the exogenous variables and leads only to other latent factors. This factor represents an intermediate effect that triggers successive behavior. This structure is given by the following matrices:

The  $\lambda$  matrix, defining the  $\eta$  latent factors in terms of the  $y^*$  discrete-choice probabilities, is specified as:

$$\lambda = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ 0 & \lambda_{22} & 0 \\ 0 & 0 & \lambda_{33} \\ 0 & 0 & 0 \end{bmatrix} \quad (14)$$

The B matrix, capturing the direct causal effects between the endogenous latent discrete-choice factors is specified to be:

$$B = \begin{bmatrix} 0 & \beta_{12} & 0 & \beta_{14} \\ 0 & 0 & 0 & \beta_{24} \\ 0 & \beta_{32} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (15)$$

This implies that  $\eta_4$ , a latent variable of arousal, directly effects both the probability of seeking highway advisory radio and the probability of diverting. In this model, similar to the previous one, the probability of seeking highway advisory radio also effects both the probability of diversion, and the probability of seeking in-vehicle navigation information.

As in the previous case, estimation is accomplished by applying the generalized least squares moment estimation method to a correlation matrix formed by tetrachoric correlations between the endogenous variables and polyserial correlations between the endogenous variables and the exogenous variables. The results are shown in Figures 3 and Table 4. This model was found significant at  $p=0.931$  and has a chi-square of 11.55 with 20 degrees of freedom. The model also has an AGFI of 0.991 and RMSR of 0.0456.

After running several model structures, each with varied matrix forms, the model presented here was found to be the only acceptable model. All t-scores were significant and there were no significant modification indices to suggest that other parameters would better the model fit. Table 5 lists other models that were tested but rejected. In each of these models there were variables in the beta and lambda- $\gamma$  matrices that had t-values that were not significant. Modification matrices revealed by LISREL also showed that these models could be improved by altering the matrix structures.

The acceptable model illustrates some interesting findings on behavior and the presence of latent arousal and motivation concepts. Eta-1, eta-2, and eta-3 are latent concepts that reflect unmeasurable effects that influence the endogenous variables. The three latent factors, each tied to a specific output, describe levels of motivation and need to modify behavior. The latent factor tied to diversion combines the effects of the arousal latent factor with the effects from the HAR-related latent factor with additional inputs from the input variables. It acts as a filter to suggest that once arousal has occurred other effects motivate drivers to divert. Increased desire to divert comes from previous diversion behavior and less familiarity of the current path.

The fourth latent variable, eta-4, is linked to other latent factors but not tied to a specific output. This can be interpreted as negative arousal potential. This variable has a negative effect on diversion and information acquisition. This factor, comprised of **enddist**, **aslink**, **VMS**, and **typet1** indicates conflict arousal as a precursor to activity. When players are further from the



destination and traveling on freeway links arousal is initiated through a marked decrease in average travel speed and a variable message indicating congestion or an incident ahead.

Similar effects occur on the two other latent factors that are tied to the information acquisition outputs. Once arousal has occurred, different variables effect whether the driver will divert or seek additional information. The negative total effects for **mapatl** and **onalink** on **HAR** and **IVNS** indicate that less familiarity with the alternate paths is more likely to lead to drivers acquiring additional information. Figure 4 illustrates the total effects on the dependent variables.

## DISCUSSION

Two models were estimated using LISREL and structural equation modeling techniques. Table 6 summarizes the model statistics. Both models fit very well and cannot be rejected at the  $p = 0.05$  level.

Modeling information acquisition as an output clearly shows that enroute behavior involves more than diversion. The act of acquiring information is itself an important behavioral response to the environment. The link between HAR and IVNS suggests that drivers may rely on several methods of information presentation. Traffic condition information and route guidance serve different but important purposes in assisting drivers to assess travel conditions and make diversion decisions.

There is an improvement when moving from the model of path analysis to the model that includes latent factors. The total effects table for the path analysis states that several inputs induce both diversion and information acquisition but no other information is known concerning the decision process. Alternatively, the model that includes latent factors provides more understanding to the behavioral choice. The four latent variables can be viewed as conflict and arousal concepts that trigger diversion and information acquisition. In turn, this arousal combined with other variables leads to different observed outcomes.

## CONCLUSIONS

This paper presented a structural equation approach to modeling enroute driver behavior. Two models were estimated, the first presenting a path analysis approach and the second involving latent factors. There are several insights to be gained from these models. A structural equation approach to model driver behavior, is useful on its own or in conjunction with traditional choice models. While structural equation models are not predictive by nature, the ability to recognize direct and indirect causal links between variables can be a valuable asset in determining impacts on behavior. Structural equations have also been used for longitudinal studies involving panel data and repeated behavior. While not used in this form in this research, this approach may lend itself to studying longer-term impacts on driving behavior.

The ability to estimate a model using latent variables enables a direct analysis of the driver and the decision making process. In the field of driver behavioral analysis, understanding the role of the driver is more critical than knowing the relationship between inputs and outputs. While category models and logit analyses provide these simple relationships, they often fail to provide an understanding of several aspects of the human element. The addition of latent factors, as shown in the models presented here, provides a deeper, more logical understanding of the decision-making process by revealing how the observed input factors are combined and used by drivers to make travel choices.

Furthermore, being able to jointly model the endogenous variables provides an alternative to nested or joint logit modeling. These models illustrated the ease with which one can model the impacts of information on diversion while having both variables as outputs. Most other modeling approaches have exclusively used information acquisition as inputs.

Although the test case for this research was a special event trip, it is likely that other types of trip could have been substituted and similar results obtained. The results on diversion behavior and information acquisition, while not surprising, confirm many basic hypotheses of travel behavior and the potential impact of ATIS.

However, it must be emphasized that the intent of case study was to demonstrate the feasibility of the prototype game and the theoretical model; generalizations regarding the population are clearly unwarranted and not made. A larger and more robust data set is needed to better test the theories. More emphasis needs to be placed on characteristics of individual behavioral characteristics and travel objectives - neither of which were included into the modeling process. A longitudinal study of repeated trips would be useful to study dynamic aspects of driver behavior.

## **ACKNOWLEDGMENTS**

The research work presented in this paper was supported by the PATH program of the California Department of Transportation, the University of California Transportation Center, and the City of Anaheim through a grant from the Federal Highway Administration. We gratefully acknowledge the support. The contents of this report reflect the views of the authors. The contents do not necessarily reflect the views of the sponsoring parties, the Institute of Transportation Studies, or the University of California.

## REFERENCES

- Adler, J. L., Recker, W. W., and McNally, M. G. (1993). In-laboratory experiments to analyze enroute driver behavior under ATIS. Presented at the 72nd Annual Meeting of the Transportation Research Board, Washington, D.C.
- Adler, J. L., Recker, W. W., and McNally, M. G. (1992a). A conflict model and interactive simulator (FASTCARS) for predicting enroute assessment and adjustment behavior in response to real-time traffic condition information. Presented at the 71st Annual Meeting of the Transportation Research Board, Washington, D.C.
- Adler, J. L., Recker, W. W., and McNally, M. G. (1992b). Using interactive simulation to model driver behavior under ATIS. Proceedings of the 4th International Conference on Microcomputers in Transportation, July 1992, Baltimore MD.
- Ben-Akiva, M., De Palma, A., and Kaysi, I. (1991). Dynamic network models and driver information systems. Transportation Research 25A, No. 5. 251-266.
- Bollen, K. A. (1989). Structural Equations With Latent Variables, John Wiley and Sons, New York.
- Browne, M.W. (1974). Generalized least squares estimators in the analysis of covariance structures. South African Statistical Journal, 8: 1-24. (Reprinted in D.J. Aigner and A.S. Goldberger (eds.), 1977), Latent Variables in Socio-Economic Models. Amsterdam: North-Holland.
- Browne, M.W. (1984). Asymptotic distribution free methods in analysis of covariance structures. British Journal of Mathematical and Statistical Psychology, 37: 62-83.
- Jöreskog, K. G. and Sörbom, D. (1989). LISREL<sup>®</sup> 7 User's Guide, Scientific Software Inc., Mooresville, IN.
- Khattak, A.J., Schofer, J.L., and Koppelman, F.S. (1993). "Commuters' Enroute Diversion and Return Decisions: Analysis and Implications for Advanced Traveler Information Systems". Transportation Research A 27A, 101-111.
- Kirk, D.D. (1973). On the numerical approximation of the bivariate normal (tetrachoric) correlation coefficient. Psychometrika, 38: 259-268.
- Olsson, U., F. Drasgow and N. J. Dorans (1982). The polyserial correlation coefficient. Psychometrika, 47: 337-347.

---

**Table 1 Case Study Variables**

**ENDOGENOUS VARIABLES**

DIVERT	0 = Did not divert, 1 = Divert
HAR	0 = Did not use HAR, 1 = Used HAR
IVNS	0 = Did not use IVNS, 1 = Used IVNS

**EXOGENOUS VARIABLES**

ENDDIST	Distance from destination - trip quarter [1-4]
VMS	Level of message sign [0-4]
TYPETL	Current path road type
TYPEATL	Alternate path road type
MTL	Familiarity with current path
MATL	Familiarity with alternate path
ONTL	Previous times on current path (past trips)
ONATL	Previous times on alternate path
ASLINK	Average link speed
SPDRATIO	Observed travel speed / Expected speed
NUMDIVER	Previous diversions during this trial

---

**Table 2 Summary of Rejected Models for Path Analysis**

Beta Matrix	X <sup>2</sup>	p	d.f.	AGFI	RMSR	PROBLEMS
(2,1), (3,2)	44.84	0.001	19	0.962	0.041	high chi-square
(2,1), (3,1)	57.14	0.000	19	0.952	0.058	not significant beta
(1,2), (2,1), (3,1), (3,2)	11.45	0.832	17	0.989	0.030	not significant beta
(1,3), (2,3)	56.04	0.000	19	0.953	0.055	not significant beta

**Table 3 Beta Matrix (t - statistics)**

	DIVERT	HAR	IVNS
DIVERT	-	-0.376	-
HAR	-	-	-
IVNS	-	0.361	-
		(3.382)	

**Gamma Matrix (t - statistics)**

	ENDDIST	VMS	TYPETL	TYPEATL	MAP	AMAP	ONTL	ONATL	ASLINK	SPDRAT	NMDIVR
DIVERT	-0.323 (-4.914)	0.228 (3.984)	-0.259 (-4.258)	-	-0.266 (-4.599)	-	-	-	-0.326 (-7.783)	-	0.758 (12.282)
Seek	-0.221	0.156	-	0.119	-	-0.283	-	-0.145	-	-0.327	-
HAR	(-2.589)	(2.846)	-	(2.311)	-	(-2.753)	-	(-2.249)	-	(-6.144)	-
Seek	-0.173	-	-	-0.335	-	-	-0.219	-	-	-	-
IVNS	(-2.565)	-	-	(-5.038)	-	-	(-1.912)	-	-	-	-

**Table 4 Beta Matrix (t - statistics)**

	Eta-1	Eta-2	Eta-3	Eta-4
Eta-1	-	-0.222 (-3.490)	-	00.283 (-4.888)
Eta-2	-	-	-	-0.153 (-2.977)
Eta-3	-	0.353 (3.331)	-	-
Eta-4	-	-	-	-

**Latent Variable Model Gamma Matrix (t - statistics)**

	ENDDIST	VMS	TYPETL	TYPEATL	MAP	AMAP	ONTL	ONATL	ASLINK	SPDRAT	NMDIVR
ETA1		-	-	-	-0.227 (-4.201)	-	-	-	-	0.130 (2.779)	0.748 (12.531)
ETA2		-	-	0.146 (3.018)	-	-0.236 (-2.070)	-	-0.144 (-2.124)	-	-0.147 (-1.914)	-
ETA3	-0.178 (-2.894)	-	-	-0.309 (-4.575)	-	-	-	-	-	-	-
ETA4	1.000	-0.737 (-3.514)	0.990 (3.782)	-	-	-	-	-	-	1.399 (4.076)	-

---

**Table 5 Summary of Rejected Models for Latent Factors Analysis**

<b>Beta Matrix</b>	<b>X<sup>2</sup></b>	<b>p</b>	<b>d.f.</b>	<b>AGFI</b>	<b>RMSR</b>	<b>PROBLEMS</b>
(1,4), (2,4), (3,4)	31.26	0.069	21	0.976	0.742	non sig. betas
(2,3), (1,4), (2,4)	22.93	0.348	21	0.982	0.058	high mod indicesB
(1,4), (2,1), (3,1)	55.45	0.000	21	0.958	0.065	high chi-square

---

---

**Table 6 Summary of Structural Equation Models**

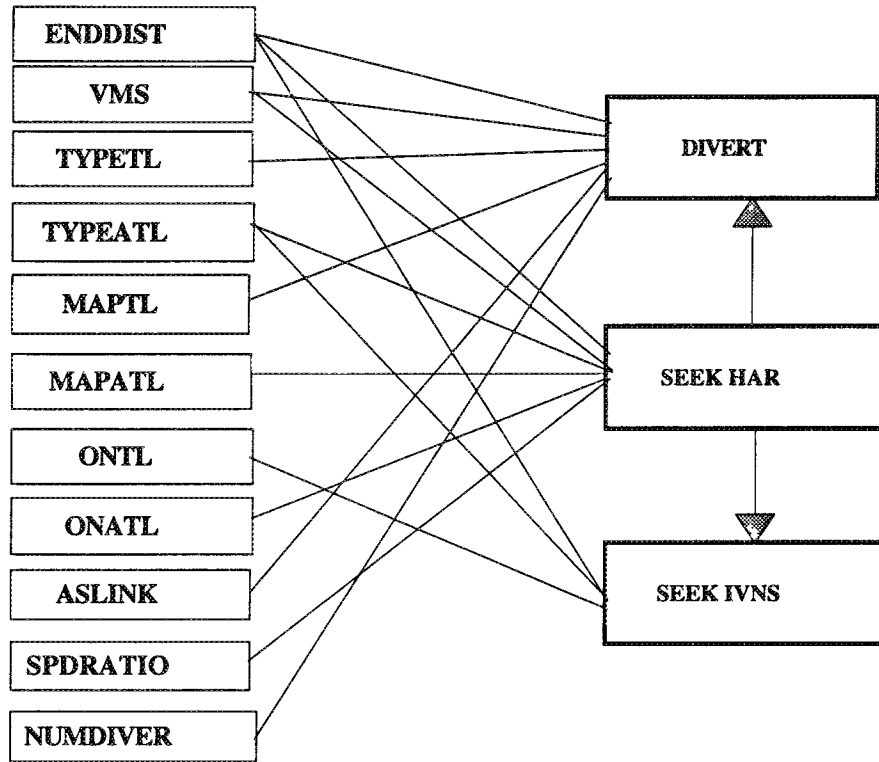
<b>Model</b>	<b>X</b>	<b>Y</b>	<b>Latent</b>	<b>X<sup>2</sup></b>	<b>p</b>	<b>d.f.</b>	<b>AGFI</b>	<b>RMSR</b>
Path Analysis	11	3	0	12.223	0.876	19	0.990	0.0303
Latent Variables	11	3	4	11.551	0.931	20	0.991	0.0456

---



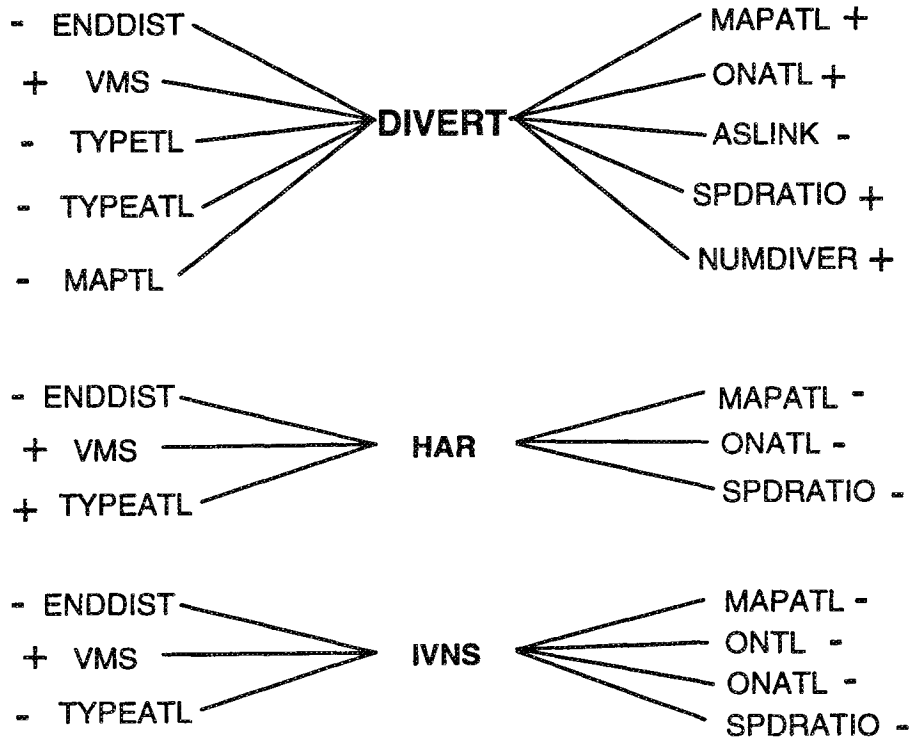
---

**Figure 1 Full Multivariate Regression on Divert and ATIS**



$X^2$  with 19 degrees of freedom = 12.223 ( $p = 0.876$ )  
AGFI = 0.990 RMSR = 0.0303

**Figure 2 Path Analysis Total Effects**

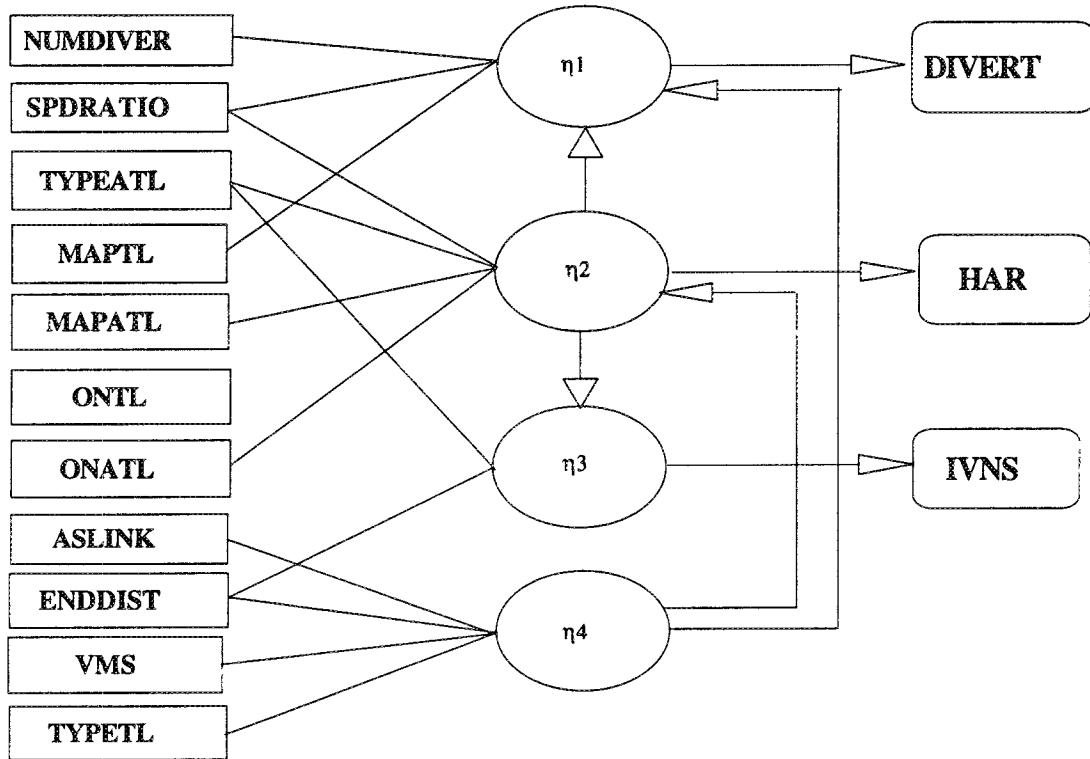


**Total effects X on Y**

	<b>enddist</b>	<b>vms</b>	<b>typetl</b>	<b>typeatl</b>	<b>maptl</b>	<b>mapatl</b>
<b>divert</b>	-.240	.170	-.259	-.045	-.266	.106
<b>har</b>	-.221	.156		.119		-.283
<b>ivns</b>	-.253	.056		-.292		-.102
	<b>ontl</b>	<b>onatl</b>	<b>aslink</b>	<b>spdratio</b>	<b>numdiver</b>	
<b>divert</b>		.055	-.326	.123	.758	
<b>har</b>		-.145		-.327		
<b>ivns</b>	-.219	-.052		-.118		

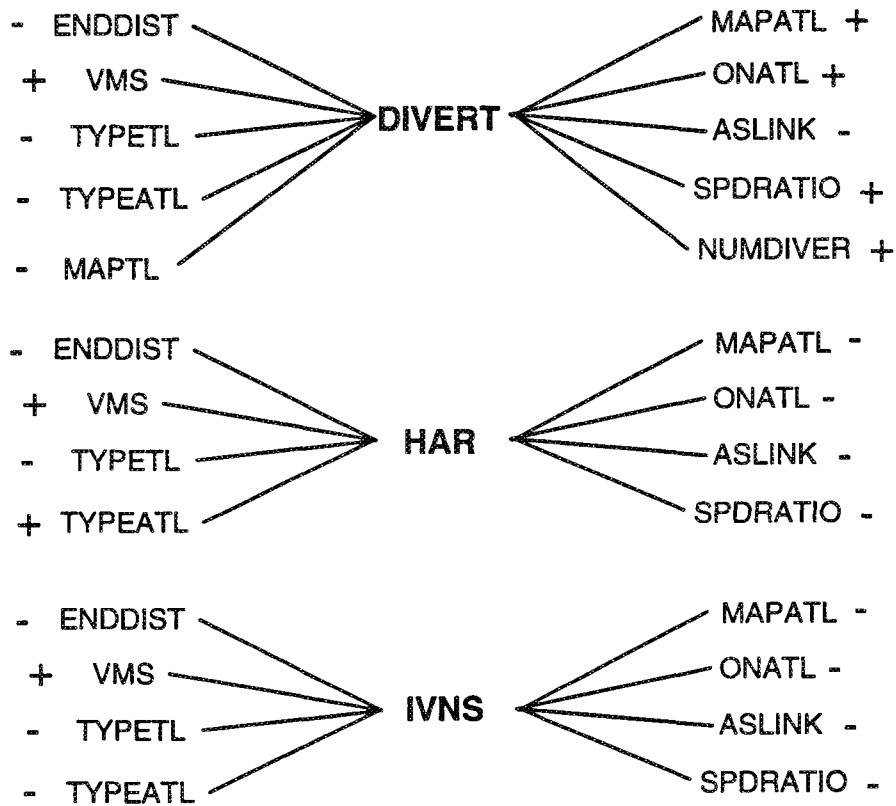
---

**Figure 3 Model with Latent Variables**



$X^2$  with 20 degrees of freedom = 11.55 ( $p = 0.931$ )  
AGFI = 0.991 RMSR = 0.0456

**Figure 4 Latent Variable Analysis Total Effects**



**Total effects X on Y**

	enddist	vms	typetl	typeatl	maptl	mapatl
<b>divert</b>	-2.49	.183	-.246	-.032	-.227	.052
<b>har</b>	-1.53	.113	-.151	.146		-.236
<b>ivns</b>	-.232	.040	-.053	-.257		-.083

	ontl	onatl	aslink	spdratio	numdiver
<b>divert</b>		.032	-.348	.162	.748
<b>har</b>		-.144	-.214	-.147	
<b>ivns</b>		-.051	-.075	-.052	