Models of Commuters' Information Use and Route Choice: Initial Results Based on Southern California Commuter Route Choice Survey

MOHAMED A. ABDEL-ATY, KENNETH M. VAUGHN, RYUICHI KITAMURA, PAUL P. JOVANIS, AND FRED L. MANNERING

A statistical analysis of commuters' route choice behavior and the influence of traffic information is presented. The analysis is based on a 1992 computer-aided telephone interview survey of Los Angeles area morning commuters. Cross tabulations were performed on the data to explore interrelationships among variables and provide a basis for subsequent model estimation. Two sets of models were estimated: bivariate probit models of whether individuals follow the same route to work every day and whether they receive traffic information (pretrip or en route) and negative binomial models of the frequency of route changes per month on the basis of pretrip and en route traffic reports. The estimation results underscore the important relationship between the use of traffic information and the propensity to change routes. In addition, important relationships are uncovered relating to the influence that commuters' socioeconomic characteristics and the level of traffic congestion they face has on traffic information use and route change frequency.

The problem of route choice for a commute trip can be defined as choosing the best route through the transportation network, in terms of some criterion or criteria, while facing temporal (i.e., departure and arrival times) and geographic (i.e., origin and destination) constraints. This best route most often is thought of as the one that minimizes travel disutility (e.g., travel time, distance, or generalized travel cost). In reality, the problem of route choice faced by an automobile driver is complex because of the large number of possible alternative routes through the road network and the complex patterns of overlap between the various route alternatives (*1*).

In recent years, an abundance of research has focused on commuters' route choice with an emphasis on how real-time traffic information might affect this choice (*2–6*). In an ongoing Partners for Advanced Transit and Highways (PATH) project at the University of California (UC) Davis entitled ATIS (Advanced Traveler Information Systems) Impact on Travel Demand, a variety of issues about traveler response to information are being investigated [see, for example, previous work (*7–9*)]. These earlier papers focused on development of learning models of drivers' adaptation to traffic advice, particularly when the advice is not always reliable. A second part of the project deals with studying the actual route choices of drivers, with the objective of developing refined route choice models that can include the effect of traveler information. Understanding route choice behavior is essential to improving network assignment methods and to investigating ATIS effectiveness (e.g., how much information drivers have or need or how information affects route choice behavior). This paper is concerned with the second part of the project.

M. A. Abdel-Aty, K. M. Vaughn, R. Kitamura, and P. P. Jovanis, Institute of Transportation Studies, 2028 Academic Surge, University of California at Davis, Davis, Calif. 95616. F.L. Mannering, Department of Civil Engineering, 121 More Hall FX-10, University of Washington, Seattle, Wash. 98915.

To probe into drivers' route choice behavior, a telephone survey of Los Angeles area morning commuters was conducted as part of the project. The survey was designed to investigate how much information drivers have about their routes, their awareness of alternate routes, their awareness of traffic conditions that could affect their route choices, and their use of available traffic information either en route or pretrip. The survey, undertaken in May and June 1992, is differentiated from those of previous studies in that the specific routes taken by individuals were obtained for their morning commute.

This paper describes the survey design and administration. General descriptive statistics are also introduced to show the characteristics, preferences, and perceptions in commuters' route choice behavior. Bivariate probit models of traffic information use and propensity to use alternative routes are also developed. In addition, negative binomial models are used to assess frequency of commuters' route changes on the basis of traffic reports, route characteristics, and individual characteristics. Further details about the survey itself and additional descriptive statistics are contained in a project report (*10*).

SURVEY AND SAMPLE DESCRIPTION

A route choice/traffic information survey was developed to target Los Angeles area morning commuters. A mail-out/mail-back survey instrument was initially designed to gather detailed information on commuters' main and alternate routes, to determine the level of information commuters have about these routes, to measure commuters' attitudes toward, and perceptions of, these routes, and to determine how existing traffic information affects their route choice behavior. The mail survey instrument required several branchings, increasing its level of complexity and potentially jeopardizing the response rate and response accuracy. Therefore, it was decided to perform a computer-aided telephone interview (CATI) survey. A CATI survey allows interviewer–respondent interaction and automatically handles branchings with complete reliability and lower interviewer error. The survey targeted a random sample of adult commuters residing in the area covered by the South Coast Air Quality Management District, which includes most of the contiguously populated areas of Los Angeles, Orange, San Bernardino, and Riverside Counties. The sampling, based on a Mitofsky-Waksberg cluster sampling design (*11*), covered both listed and unlisted numbers. The Mitofsky-Waksberg sampling is known to reduce the number of unproductive dialings and improves efficiency (*12*). In all, 944 commuters were surveyed in May and early June 1992. Summary statistics for the sample are presented in Table 1. The values shown in this table are within expected bounds.

To test the representativeness of the sample, several socioeconomic and commute characteristics were compared and statistically tested with the 1990 Census (*13*), the 1991 California Statewide Travel Survey results (CSTS) (*14*), and the 1990 California Statistical Abstract (*15*). In most cases the null hypothesis that the values from the route choice survey are not different from the corresponding statistical sources was not rejected at the 0.05 level of significance, implying that the sample well represents the population in the study area. Further information on this and other tests can be found elsewhere (*10*).

Traffic Information Use

As Table 1 shows, the survey provided some interesting insight into travelers' use of traffic information and their choice of route. In the survey, traffic information questions were divided into two groups, depending on where the information is received, either before (pretrip) or while (en-route) driving to work. About 36.5 percent of the respondents listen to traffic reports before leaving their homes, and 51.25 percent listen while driving. Close to 27.6 percent of the respondents listen to traffic reports both at home and en route, and 60.1 percent listen to reports either at home or en route, whereas 39.9 percent never listen to reports. These findings are consistent to a great extent with those of Khattak et al. (*2*). Most respondents who receive traffic information perceive traffic reports to be either very accurate or somewhat accurate.

More women (40 percent) listen to traffic reports before leaving home to work than men (33 percent), whereas more men (54.5 percent) listen to reports en route than women (47.7 percent). The hypothesis of no differences between sexes was rejected using Pearson chi-square at a 0.05 level of significance. It was also found that more women change their routes or departure times as a result of listening to traffic reports before leaving their homes, whereas men change their routes more frequently than women as a result of traffic reports they hear while driving to work. However, it is possible that socioeconomic or commute characteristics, or both, associated with gender are the cause of such differences between men and women and not gender itself.

Commuters who use freeways may be more likely to receive traffic information if their freeway traffic conditions are perceived as heavy or very heavy. The relationship was confirmed (using a chi-square test) for pretrip information but not for en-route information. This suggests that commuters try to find out their freeway conditions in advance, possibly because these are the segments of their routes that are exposed most to delays or because they realize that their diversion options, once they get onto a freeway, are very limited, or both.

Route Choice Behavior

Only 15.5 percent of the respondents said they use more than one route to work. Considering the well-developed freeway network in the study area, this may be considered a low percentage. However it indicates that an information system that would make people aware of alternative routes has promising potential. About 50 percent of the respondents had at least one freeway segment in their primary routes (a primary route is the route that the respondent uses most frequently), and 38 percent had at least one freeway segment in their secondary routes (Figure 1); secondary routes tend to have more surface streets than primary routes, possibly as alternatives used to avoid congestion on freeways. The percent of freeway users in the CSTS data is 46.3 percent, which is very close to the results of the present study. Even for an area that is generally considered saturated with freeways, 50 percent of the primary routes involves no freeway at all.

Finally, it is interesting that the most frequent reason for changing routes, cited by 34 percent of respondents, is the traffic that the respondents see on the roads. The need to make stops on the way

TABLE 1 Sample Summary Statistics

• Commute distance on usual route (miles)	12.75
• Travel time on usual route (minutes)	28.14
• Trip duration (including stops)	31.9
• Percent of respondents commuting in single-occupant autos/carpool/public transit	78.8/14.6/4.9
• Percent receiving pre-trip traffic reports	36.5
• Percent receiving en-route traffic reports	51.25
• Percent of respondents with flexible/ somewhat flexible / fixed work starting time	24.4/30.4/45.2
• Percent male/female	51.3/48.7
• No. of household cars	2.31
• No. of years at present address	7.24
• No. of years at present job location	5.52
• Percent own/rent their homes	59/41
• Household income	38,750
• Percent of college graduates	43.8
• Think traffic congestion is a problem or major problem (percent)	61.3
• Think trip time uncertainty is a problem or major problem (percent)	31.9

Note: Values are averages unless noted otherwise.

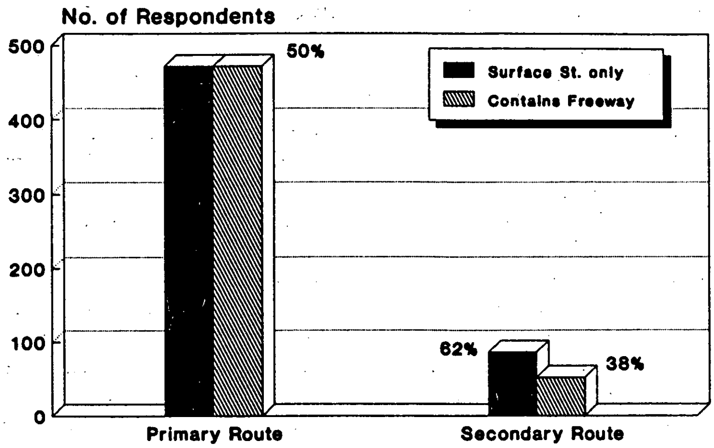

FIGURE 1 Composition of primary and secondary routes by road type.

and traffic reports follow (15.5 and 14 percent, respectively). Additional reasons include the time of day (8 percent) and the day of the week (5.5 percent). If the percent of respondents that base their choice on the traffic they see is added to others who base their choice on traffic reports, then about 50 percent of the commuters depend on real-time information for choosing their routes. This finding reiterates the potential of an ATIS system to alleviate traffic delays.

Individuals with higher incomes tend to report using more than one route to work. The fraction of individuals with alternative routes (percent of multiple route users within each income category) increases from 6.7 percent among those with incomes less than \$25,000 to 28 percent among those with incomes more than \$100,000. The null hypothesis of independence between income and using alternative routes is rejected. Khattak et al. (*2*) also found that higher-income drivers were more likely to take alternate routes. The same relationship is also found for level of education: highly educated people tend to use alternate routes.

MODELING APPROACH

To assess commuters' propensity to change routes and acquire traffic information, the study focused on the joint decision of whether commuters follow the same route to work every day and whether they receive traffic information (pretrip or en route). The objective is to examine the association between information use and route choice and to verify the results of the cross-tabulation analysis in multivariate modeling contexts. For such a joint decision, the bivariate (two-dimensional) probit formulation is appropriate. Commuters' frequency of route changes on the basis of traffic information is then modeled using negative binomial regression models. The modeling effort reported in this paper represents a preliminary analysis of the interplay of information use and route choice. The variables considered in model development include the attributes of main commute routes, attributes of commuters, and their perception of traffic conditions. Future work could extend the range of variables to include objectively measured traffic characteristics for the respective commuters' main and alternative routes. Figure 2 summarizes the modeling effort presented.

Joint Estimation of Route Switching and Information Choices

There is a need to identify the factors that lead a commuter to use single or multiple routes to work and to receive traffic information. Gaining an understanding of this issue will aid in how traffic conditions and other factors affect the use of traffic information and route switching. In particular, building a model that predicts route-switching behavior as a function of information use will aid in evaluating the potential impacts of ATIS on route choice.

Methodological Approach

The commuters' choice of receiving traffic information and their use of alternate routes are likely to be interrelated. As such, there is a likely correlation of unobserved effects (between information use and route choice) which if not accounted for, would lead to biased model coefficient estimates. An example of such unobserved correlation would be the tendency of a commuter to be "adventurous" and "dynamic." Clearly such a tendency would be difficult to quantify (and therefore likely to show up in model error terms), but adventurous and dynamic commuters would be expected to be much more likely to receive traffic information and to change routes. This would produce a positive correlation in error terms that must be accounted for. An appropriate model for capturing this correlation is the bivariate probit.

The bivariate probit model can be used directly to identify the contributing factors that influence route switching behavior and affect the likelihood of receiving traffic information. In this case, the two choices are (*a*) whether the respondent receives traffic information ($Y_1 = 0$ or 1), and (*b*) whether the respondent uses more than one route to work ($Y_2 = 0$ or 1). These two choices can be represented by the following simultaneous equation system:

$$Y_1^* = \beta X_1 + \epsilon$$

$$Y_1 = \begin{cases} 1 & \text{if } Y_1^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

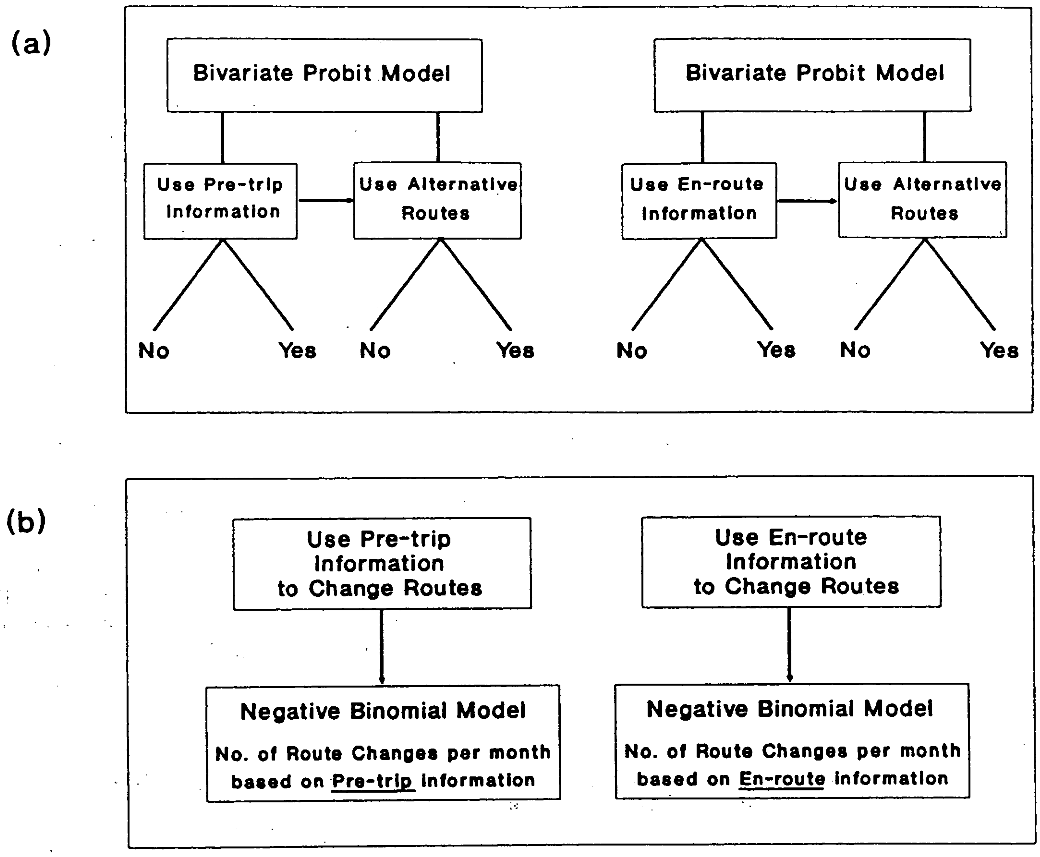

FIGURE 2 Modeling structure.

$$Y_2^* = \alpha X_2 + \Theta Y_1 + \xi$$

$$Y_2 = \begin{cases} 1 & \text{if } Y_2^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

where

- Y_1^* = latent variable indicating the respondent's propensity to listen to traffic information;
- Y_1 = observed choice (1 if the respondent listens to information and 0 otherwise);
- Y_2^* = latent variable indicating the respondent's propensity to use multiple routes;
- Y_2 = observed choice (1 if the respondent is a multiple route user, and 0 if exactly one route is used every day to work);
- β, α = coefficient vectors;
- Θ = scalar coefficient;
- X_1, X_2 = vectors of explanatory variables influencing choice behavior; and
- ϵ, ξ = random error terms.

Assuming that ϵ and ξ are correlated $[E(\epsilon\xi) \neq 0]$, then the two equations should be estimated simultaneously using the full-information maximum likelihood (FIML) or sequentially using the limited-information maximum likelihood (*16,17*). If a limited-information approach is adopted, parameters are estimated in one equation at a time with instrumental variables (*10*) or correction terms (*18*) introduced to account for error correlation. For a linear system, these techniques provide consistent but inefficient estimates of parameters (*16*). However, in a system of two binary-choice equations, as is the case in this study, these approaches may lead to inconsistent estimates [numerical comparisons of alternative estimators are given by Kitamura (*19*)]. The FIML is desirable because it offers consistent and efficient estimates while accounting for possible error correlation across equations. Thus, FIML is adopted in this study.

Distributional assumptions need to be made on the random error terms ϵ and ξ to express response probabilities. A probit model offers a theoretically sound formulation for discrete responses. Adoption of the probit formulation in a situation involving two binary-choice endogenous variables would imply that the joint distribution of ϵ and ξ is given by the bivariate standard normal distribution, with var(ϵ) = var(ξ = 1 for normalization.

For this system of equations (i.e., Equations 1 and 2), the full-information likelihood function for the bivariate probit is developed by first defining sample strata as follows:

S_1: $Y_1 = 1$ $\quad$ $Y_2 = 1$

S_2: $Y_1 = 1$ $\quad$ $Y_2 = 0$

S_3: $Y_1 = 0$ $\quad$ $Y_2 = 1$

S_4: $Y_1 = 0$ $\quad$ $Y_2 = 0$

The likelihood function for the first set of observations, S_1, is derived by considering the joint probability of the event, $Y_1 = 1$ and $Y_2 = 1$:

$$Pr[Y_1 = 1, Y_2 = 1] = Pr[Y_1^* \geq 0, Y_2^* \geq 0]$$
$$= Pr[\epsilon \geq -\beta X_1, \xi \geq -\alpha X_2 - \theta]$$
$$= \int_{-\beta X_1}^{\infty} \int_{-\alpha X_2 - \theta}^{\infty} f(\epsilon, \xi)\, d\epsilon\, d\xi$$

where f is the standard bivariate normal density function:

$$f = \frac{2\pi}{\sqrt{(1-\rho^2)}} \exp\left[-\frac{(\epsilon^2 - 2\rho\epsilon\xi + \xi^2)}{2(1-\rho^2)}\right] \quad (3)$$

and ρ is the correlation coefficient between ϵ and ξ.

The likelihood function for this set of observations is

$$L_1 = \prod_{S_1} \int_{-\beta X_1}^{\infty} \int_{-\alpha X_2 - \theta}^{\infty} f(\epsilon, \xi)\, d\epsilon\, d\xi \quad (4)$$

Similarly L_2, L_3, and L_4 can be derived. Therefore, the likelihood function for the entire sample is

$$L = \prod_{S_1} \int_{-\beta X_1}^{\infty} \int_{-\alpha X_2 - \theta}^{\infty} f(\epsilon, \xi) d\epsilon\, d\xi$$
$$\prod_{S_2} \int_{-\beta X_1}^{\infty} \int_{-\infty}^{-\alpha X_2 - \theta} f(\epsilon, \xi)\, d\epsilon\, d\xi$$
$$\prod_{S_3} \int_{-\infty}^{-\beta X_1} \int_{-\alpha X_2}^{\infty} f(\epsilon, \xi)\, d\epsilon\, d\xi$$
$$\prod_{S_4} \int_{-\infty}^{-\beta X_1} \int_{-\infty}^{-\alpha X_2} f(\epsilon, \xi) d\epsilon\, d\xi \quad (5)$$

Parameter vectors β, α, θ, and ρ are estimated so as to maximize L. The statistical significance of the coefficient θ will indicate whether state dependence is present. Also, significant error correlation between ϵ and ξ (ρ) will indicate the presence of unobserved individual factors (heterogeneity) that affect both choices of route and receiving information.

Estimation Results for the Bivariate Probit Models

Two bivariate probit models were developed after investigating several alternative model formulations. The first estimates whether respondents often receive traffic reports before leaving home to work (pretrip) and whether they are multiple-route users. The second estimates whether respondents often receive traffic reports while driving to work (en route) and whether they are multiple-route users (the whole sample is used in estimating these models).

Estimation results for the pretrip information/multiple-route user model are given in Table 2. All variables included are self-explanatory and their coefficients are readily interpretable. The pretrip information model indicates that people who perceive no variation in traffic conditions on their usual commute route are less likely to listen to pretrip traffic reports. Women, long-distance commuters, or respondents who reported uncertainty in travel time as a major problem, or all of these, are more likely to listen to these

TABLE 2 Bivariate Probit Model Estimating Whether Respondents Receive Traffic Reports Before Leaving Home to Work and Whether They Are Multiple-Route Users

	Coefficient	t-statistic
PRE-TRIP INFORMATION MODEL		
Constant	-0.416	-3.79
X_1 No variation in traffic conditions dummy (1 if no variation is perceived, 0 otherwise)	-0.361	-3.68
X_2 Female dummy (1 if female, 0 otherwise)	0.110	1.15
X_3 Uncertainty of travel time dummy (1 if reported that trip time uncertainty is a major problem, 0 otherwise)	0.436	3.23
X_4 Distance from home to work	0.013	3.40
MULTIPLE ROUTE MODEL		
Constant	-2.033	-6.95
X_5 Income dummy (1 if income ≥ $75,000, 0 otherwise)	0.302	2.43
Y_1 Receiving pre-trip information dummy (1 if receive pre-trip information, 0 otherwise)	1.002	2.74
X_6 No. of driving days in the last 2 weeks	0.032	1.26
X_7 Level of education dummy (1 if respondent is a college grad. or completed some college, 0 otherwise)	0.409	2.55
Error-term Correlation	-0.518	-2.38

Summary Statistics
Log Likelihood at zero = -1061.761
Log Likelihood at market share = -790.804
Log Likelihood at convergence = -758.191
Likelihood ratio index = 0.286
Number of observations = 733
Percent correct predicted = 72%

Note: Variables' coefficients are defined for receiving reports and multiple route use

TABLE 3 Bivariate Probit Model Estimating Whether Respondents Receive Traffic Reports While Driving to Work and Whether They are Multiple-Route Users

	Coefficient	t-statistic
EN-ROUTE INFORMATION MODEL		
Constant	-0.303	-2.82
X_1 No variation in traffic conditions dummy (1 if no variation is perceived, 0 otherwise)	-0.244	-2.42
X_2 College graduate dummy (1 if respondent is a college grad, 0 otherwise)	0.195	2.00
X_3 Uncertainty of travel time dummy (1 if reported that trip time uncertainty is a major problem, 0 otherwise)	0.708	4.51
X_4 Distance from home to work	0.026	6.57
MULTIPLE ROUTES MODEL		
Constant	-2.061	-5.92
X_5 Income dummy (1 if income ≥ $75,000, 0 otherwise)	0.306	2.33
Y_1 Receiving en-route information (1 if receive en-route information, 0 otherwise)	0.531	1.66
X_6 No. of driving days in the last 2 weeks	0.035	1.28
X_7 Level of education dummy (1 if respondent is a college grad. or completed some college, 0 otherwise)	0.415	2.44
Error-term Correlation	-0.174	-0.82

Summary Statistics
Log Likelihood at zero = -1061.761
Log Likelihood at market share = -815.902
Log Likelihood at convergence = -762.335
Likelihood ratio index = 0.282
Number of observations = 733
Percent correct predicted = 84.9%

Note: Variables' coefficients are defined for receiving reports and multiple route use

reports. These findings are consistent with those of an earlier analysis of the data.

For the multiple-route choice model, high income (≥ $75,000), high level of education (college graduate or completed some college), and the number of days driving to work in 2 weeks increase the likelihood of using multiple routes. The positive coefficient of receiving pretrip information indicates that commuters who receive this information are more likely to use more than one route to work, whereas the significance of the variable indicates the important effect of Y_1^* on Y_2^*. The significance of the correlation between the two error terms underscores the importance of accounting for cross-equation correlation. The negative sign indicates the presence of unobserved factors that reversely affect the two behavioral aspects. Note that there is no expectation with respect to the sign of the error correlation. For example, a cautious and well-prepared commuter would try to obtain as much information as possible before departure (positive ϵ) but choose to adjust the departure time rather than venture onto unfamiliar alternate route (negative ξ).

Estimation results for the en-route information/multiple-route user model are given in Table 3. The model is similar to the previous model, except that gender is replaced by a college graduate dummy, which significantly increases the likelihood that a respondent receives en-route traffic reports. The positive coefficient of receiving en-route information (Y_1) indicates that commuters who receive en-route information are more likely to use more than one route to work, although this variable is not highly significant (only at the 90 percent level of significance). The correlation between the two error terms is insignificant.

Frequency of Changing Routes On the Basis of Information

To assess commuter frequency in changing routes on the basis of traffic information, an appropriate statistical modeling technique is needed. The Poisson regression was initially attempted, but chi-square tests indicated significant differences between the estimated and observed route switching frequencies. Also, the Poisson distribution was rejected because the mean and variance of the dependent variables are different, indicating substantial overdispersion in the data (number of route changes based on pretrip information: mean, 1.69, variance, 7.34; number of route changes based on en route information: mean, 1.46, variance, 6.61). Such overdispersion suggests a negative binomial model. The negative binomial model is an extension of the Poisson regression model and allows the variance of the process to differ from the mean.

Methodological Approach

This section on methodological approach is drawn from Greene (*20*). The negative binomial model arises from the Poisson model by specifying

$$\ln \lambda_i = \beta X_i + \epsilon \quad (6)$$

where

- λ_i = parameter giving individual i's expected route-changing frequency;
- β = vector of estimable parameters;
- X_i = vector of commuting and socioeconomic characteristics for individual i; and
- ϵ = error term, where $\exp(\epsilon)$ has a gamma distribution with mean 1 and variance α^2. The resulting probability distribution is as follows:

$$Pr[Y = y_i \mid \epsilon] = \exp[-\lambda_i \exp(\epsilon)]\lambda_i^{y_i}/y_i! \quad (7)$$

where y_i is the number of route changes, and all other variables are as previously defined. Integrating ϵ out of this expression produces the unconditional distribution of y_i. The formulation of this distribution is

$$Pr[Y = y_i] = \Gamma(\theta + y_i)/[\Gamma(\theta)y_i!]\, u_i^{\theta}(1 - u_i)^{y_i} \quad (8)$$

where

- $P[Y = y_i]$ = probability of commuter i making y_i changes in a specified period of time.
- $u_i = \theta/(\theta + \lambda_i)$
- $\theta = 1/\alpha$

Compared with the Poisson model, this model has an additional parameter α, such that

$$\text{Var}[y_i] = E[y_i]\{1 + \alpha E[y_i]\} \quad (9)$$

This is a natural form of overdispersion in that the overdispersion rate is

$$\text{Var}[y_i]/E[y_i = 1 + \alpha E[y_i] \quad (10)$$

Such an approach is well suited to modeling frequency of route change because it accounts for the no-change option ($y_i = 0$) as well as all other possible non-negative integer outcomes (*21*). The negative binomial model can be estimated by standard maximum likelihood methods.

Testing for Existence of Selectivity Bias

Before proceeding with the estimation of the negative binomial models, it is important to test for possible selectivity bias. Selectivity bias could be present if the commuters observed to be using traffic information as a basis for changing routes were a self-selected group with route change behavior that systematically differed from those commuters not observed to be using information as a basis for changing routes. Such selectivity creates a problem because frequency data have been collected only on those individuals observed to be using information for changing routes. If their behavior systematically defers from those not observed changing routes, the estimates of β will be biased.

Selectivity bias correction methods in standard regression equations have been derived by other researchers (*22,23*). However, developing corrective techniques for count data (i.e., on the basis of a negative binomial regression) has not been done and is likely to be a difficult task because a closed-form expression for the expected value of the gamma error term (see Equation 6) conditioned on the bivariate probit error terms must be developed [i.e., E $(\epsilon \mid \epsilon\xi)$]. Such a formulation is beyond the scope of this paper. However, a suggestive test of this matter using a standard discrete–continuous selectivity bias correction procedure (*24*) was conducted. In doing so, the bivariate probit model (of whether or not information is used) with a simple independent binary logit model and the negative binomial regression model (of the frequency of route changes) with a standard regression model were approximated.

Formalizing this, the utility, to respondent i, of using traffic information, U_{iI}, can be written as

$$U_{iI} = \beta_1 X_1 + \epsilon_i \quad (11)$$

where

- β_1 = vector of estimable parameters,
- X_1 = vector of factors influencing information use, and
- ϵ_i = Gumbel distributed error term.

These variables give rise to the binary logit formulation

$$P_{iI} = 1/[1 + \exp(-\beta_1 X_1)] \quad (12)$$

where P_{iI} is the probability of respondent i using information.

For the regression equation of the frequency of route changes conditioned on the use of information, I, the following is given:

$$E(y_i \mid I) = \beta_2 X_2 + E(\psi_i \mid I) \quad (13)$$

where

- β_2 = vector of estimable parameters,
- X_2 = vector of factors influencing the frequency of route choice (y_i), and
- ψ_i = a normally distributed error term.

Selectivity bias arises because of unobserved effects [i.e., $E(\epsilon_i\psi_i) \neq 0$]. To correct this bias in the estimation of route change frequency, an expression for the conditional expectation of the error term [i.e., $E(\psi_i \mid I)$] is needed. Dubin and McFadden (*23*) have shown this to be

$$E(\psi_i \mid I) = \rho(\sigma_\psi/\sigma_\epsilon)\, \pi_{iI} \quad (14)$$

where

- σ_ψ = standard deviation of the normally distributed error term ψ_i,
- σ_ϵ = standard deviation of the logistic error term (from Equation 12),
- ρ = partial correction coefficient for ψ and ϵ,
- $\pi_{iI} = [1/P_{iI}][P_{iI}\log(P_{iI}) + (1 - P_{iI})\log(1 - P_{iI})]$, and
- P_{iI} = probability of respondent i choosing to use information I. Thus Equation 13 becomes

$$E(y_i \mid I) = \beta_2 X_2 + \omega\pi_{iI} \quad (15)$$

where ω is an estimable coefficient equal to $\rho(\sigma_\psi/\sigma_\epsilon)$. The significance of the coefficient, ω, associated with the selectivity correction term, π_{il}, gives a measure of the importance of selectivity bias in the equation.

The ω coefficient terms in both regression models (number of times per month changing routes on the basis of pretrip information and number of times per month changing routes on the basis of en route information) were statistically insignificant (pretrip information model $\omega = -0.001$, t-statistic $= -0.795$; en route information model $\omega = -0.001$, t-statistic $= -1.280$), suggesting that selectivity bias is not present. It is therefore concluded that estimating the negative binomial models without possible error correlation between the bivariate probit and the negative binomial is not likely to be a significant source of error.

Estimation Results for Negative Binomial Models

Two negative binomial models were developed: the first modeled the number of route changes per month on the basis of listening to pretrip traffic reports, and the second modeled the number of route changes per month on the basis of listening to en-route traffic reports. The estimation results for the first model are illustrated in Table 4. The results show that commuter's perceptions have an important effect on the number of route changes; that is, if respondents perceive substantial variation in traffic conditions from day to day on their primary route, they are likely to make more route changes per month. If it is perceived to be accurate, information will have a positive effect on the number of changes per month; dummy variables representing individual report values (e.g., 1, 2, 3) were attempted but the results showed that the relationship was linear, and therefore, a simple ordering of responses was used.

Turning to the socioeconomic factors, a high level of education (e.g, college graduates) was found to have a positive impact on the number of route changes per month. Also, from the commute characteristics, the log of travel time on the most frequently used route has a positive impact on the number of route changes per month, indicating that longer commutes make travelers more likely to change routes. A possible explanation can be that time-consuming commutes lead to a greater awareness and use of alternate routes, on the basis of pretrip information. However, the log transformation indicates that this effect diminishes with increasing travel time. Finally, the significance of the overdispersion parameter (α) indicates that the negative binomial formulation is preferred to the more restrictive Poisson formulation.

The second model (the frequency of route changes per month on the basis of en-route information) is presented in Table 5. The results show that the carpool dummy has a positive effect on the number of route changes per month on the basis of en-route traffic reports. It appears that once carpoolers are together on the road, en-route information influences their decision to change routes. A perception of substantial traffic variation and bad traffic conditions on the usual route increased the frequency of route changes. Also, the perception that information is accurate has a positive effect (again, dummy variables representing report accuracy and traffic conditions were attempted but a linear relationship was found; therefore, the ordered responses were used). Individuals' perception of reality is important because it ultimately drives their behavior, which indicates that accurate traffic information is vital for commuters who perceive variations or bad traffic conditions on changing routes.

The model also shows that freeway users tend to change routes more frequently on the basis of en-route information, possibly as a means to avoid congestion. The positive coefficient of the log of commute distance depicts that longer distances cause route changes on the basis of en-route information. The use of the log transformation indicates that this effect is nonlinear, with marginal increases in distance playing a stronger role in shorter commutes. Again, the significance of the overdispersion parameter (α) shows that the negative binomial formulation is a preferred specification.

TABLE 4 Negative Binomial Model: Frequency of Route Changes Per Month to Work On the Basis of Pretrip Reports

	ß	t-statistic
β_0 Constant	-2.735	-2.43
X_1 Perceived Variation in traffic conditions dummy (1 if traffic conditions are substantially different from day to day on the usual commute route, 0 otherwise)	0.752	1.50
X_2 Perceived accuracy of traffic reports (1 not at all accurate, 2 not very accurate, 3 somewhat accurate, 4 very accurate, 5 extremely accurate)	0.362	2.42
X_3 College graduate dummy (1 if college graduate, 0 otherwise)	0.354	1.48
X_4 Log driving time on last trip using the usual route	0.507	2.19
α overdispersion parameter	2.065	5.90

Summary Statistics
Log Likelihood at zero = -833.809
Log Likelihood at convergence = -415.779
ρ^2 = 0.501
Number of observations = 238

TABLE 5 Negative Binomial Model: Frequency of Route Changes Per Month to Work On the Basis of En-Route Reports

	ß	t-statistic
β_0 Constant	-2.385	-4.59
X_1 Carpool dummy (1 if mode is carpool, 0 otherwise)	0.473	1.68
X_2 Perceived Variation in traffic conditions dummy (1 if traffic conditions are substantially different from day to day on the usual commute route, 0 otherwise)	0.617	1.61
X_3 Rating of traffic conditions (1 very good, 2 good, 3 OK, 4 bad, 5 very bad)	0.250	2.66
X_4 Perceived accuracy of traffic reports (1 not at all accurate, 2 not very accurate, 3 somewhat accurate, 4 very accurate, 5 extremely accurate)	0.297	2.91
X_5 Freeway user dummy (1 if uses fwy, 0 otherwise)	0.420	1.56
X_6 Log commute distance	0.190	1.40
α overdispersion parameter	2.149	7.97

Summary Statistics
Log Likelihood at zero = -1426.647
Log Likelihood at convergence = -675.750
ρ^2 = 0.526
Number of observations = 443

SUMMARY AND CONCLUSIONS

This paper uses a CATI survey carried out as part of a research project at UC Davis. This survey was designed to gain a basic understanding of drivers' route choice behavior, to collect detailed information about their commute routes, and to explore how commuters use traffic information to decide on what routes to travel to work.

An analysis using general descriptive statistics showed several tendencies in the commuters' route choice decisions. Only 15.5 percent of the respondents reported that they do not always follow the same exact route to work, which indicates a potential benefit from an information system that would make more commuters aware of alternative routes.

The following were cited as the most common reasons for changing from a primary route: the desire to decrease the trip time, receiving traffic reports, and the time the commuters leave their homes. High income and a high level of education were two sociodemographic factors correlated with the use of more than one route. Other factors, such as the commute distance, did not seem to have a significant effect on using alternative routes.

Finally, the statistical exploration of the data also indicated that gender influences the use of traffic information. Women tend to listen to pretrip traffic reports more frequently than men and tend to use freeways less frequently than men.

Bivariate probit models were developed to determine the factors that influence information use and the propensity to use alternative routes. The models showed the significant influence of income, education, frequency of driving to work, and listening to traffic reports on the commuters' route choice. Also, perceived variation in traffic conditions, gender, commute distance, and travel time uncertainty affected the likelihood of listening to traffic information.

Negative binomial models were developed to assess commuters' frequency in changing routes. Two models were developed: the first modeled the number of route changes per month on the basis of pretrip traffic reports and the second modeled the number of route changes per month on the basis of en-route traffic reports. The models showed the significant effect that commuters' perceptions of the accuracy of traffic reports and variation in traffic conditions, travel time, and the level of education had on the frequency of changing routes on the basis of pretrip information. Also, traffic conditions, perceptions of information accuracy and traffic variation, freeway use, commute distance and carpool, were among the variables influencing the frequency of route changes on the basis of en-route traffic information.

The findings of this study suggest at least two important directions for future research. The first direction is methodological in nature. There is a need to develop an FIML procedure for estimating simultaneously the commuter's choice to use information and the frequency of route changes. This task will not be easy because of the complexity of the error term structure, but there are potentially many applications of such a procedure to the analysis of route choice behavior and other ATIS-related concerns.

The second direction relates to the need for information on the specific routes used by travelers. Such information would include highway geometrics, signal timings, and the temporal distribution of traffic. Although it is often tedious and time consuming to process, this information can be gathered and used to explore many detailed relationships that will have a direct impact on ATIS utilization.

ACKNOWLEDGMENTS

The authors wish to acknowledge the comments and suggestions of four anonymous referees. Their recommendations resulted in a substantially improved paper. They also thank Thomas Golob, of the University of California at Irvine, for his significant contribution to the design of the survey. Special thanks to the California Department of Transportation (Caltrans) and PATH for funding this research. The models presented in this paper were estimated using LIMDEP econometric software.

REFERENCES

1. Antonisse R., A. Daly, and M. Ben-Akiva. A Highway Assignment Method Based on Behavioral Models of Car Drivers' Route Choice. In *Transportation Research Record 1220,* TRB, National Research Council, Washington, D.C., 1989.
2. Khattak, A., J. Schofer, and F. Koppelman. Effect of Traffic Reports on Commuters' Route and Departure Time Changes. *Proc., Vehicle Navigation and Information Systems Conference,* Part 2, Society of Automotive Engineers, Warrendale, Pa., Oct. 1991, pp. 669–679.
3. Hatcher, G. and H. Mahmassani. Daily Variability of Route and Trip Scheduling Decisions for the Evening Commute. In *Transportation Research Record 1357,* TRB, National Research Council, Washington, D.C., 1992.
4. Haselkorn, M., J. Spyridakis, and W. Barfield. *Improving Motorist Information Systems: Towards a User-Based Motorist Information System for the Puget Sound Area.* Final Report. Washington State Transportation Center, University of Washington, Seattle, March 1990.
5. Haselkorn, M., J. Spyridakis, and W. Barfield. Surveying Commuters to Obtain Functional Requirements for the Design of a Graphic-Based Traffic Information System. *Proc., Vehicle Navigation and Information Systems Conference,* Part 2, Society of Automotive Engineers, Warrendale, Pa., Oct. 1991, pp. 1041–1044.
6. Mannering, F., S-G. Kim, W. Barfield, and L. Ng. Statistical Analysis of Commuters' Route, Mode, and Departure Time Flexibility. *Transportation Research,* Vol. 2C, No. 1, 1994, pp. 35–47.
7. Vaughn, K., M. Abdel-Aty, R. Kitamura, P. Jovanis, and H. Yang. Experimental Analysis and Modeling of Sequential Route Choice Behavior Under ATIS in a Simplistic Traffic Network. In *Transportation Research Record 1408,* 1993, pp. 75–82.
8. Yang, H., R. Kitamura, P. Jovanis, K. Vaughn, and M. Abdel-Aty. Exploration of Route Choice Behavior with Advanced Traveler Information using Neural Network Concepts. *Transportation* 20, 1993, pp. 199–223.
9. Abdel-Aty, M., K. Vaughn, P. Jovanis, and R. Kitamura. Understanding Traveler Responses to ATIS. *Proc., 26th ISATA Conference,* Aachen, Germany, Sept. 1993.
10. Abdel-Aty, M., K. Vaughn, R. Kitamura, and P. Jovanis. *Survey of Route Choice Behavior: Empirical Results From Southern California and Their Implications for Advanced Traveler Information Systems.* Research Report UCD-ITS-RR-93-12. Institute of Transportation Studies, University of California at Davis, 1993.
11. Groves, R., P. Biemer, L. Lyberg, J. Massey, W. Nicolls, and J. Waksberg. *Telephone Survey Methodology,* John Wiley and Sons, New York, 1988.
12. Frey, J. *Survey Research By Telephone,* Sage Publications, Beverly Hills, 1983.
13. Bureau of the Census. *1990 Census of Population and Housing, California.* U.S. Department of Commerce, Economics and Statistics Administration, Washington D.C., June 1992.
14. *1991 Statewide Travel Survey: Summary of Findings,* California Department of Transportation, Sacramento, Nov. 1992.
15. *California Statistical Abstract 1990,* Department of Finance, Sacramento, California, Oct. 1990.
16. Maddala, G. *Limited-Dependent and Qualitative Variables in Econometrics,* Cambridge University Press, Cambridge, England, 1983.
17. Kitamura, R., R. Pendyala, and K. Goulias. Weighting Methods for Choice Based Panels with Correlated Attrition and Initial Choice. In *Transportation and Traffic Theory* (C.F. Daganzo, ed.), Elsevier Science Publishers, The Netherlands, 1993, pp. 275–294.
18. Heckman, J. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement,* Vol. 5, 1976, pp. 475–492.
19. Kitamura, R. A Review of Dynamic Vehicle Holdings Models and a Proposal for a Vehicle Transaction Model. *Proc., Japanese Society of Civil Engineers,* 440/IV-16, 1992.
20. Greene, W. *LIMDEP: User's Manual and Reference Guide, Version 6.0.* Econometric Software, Inc., Bellport, N.Y., 1992.
21. Mannering, F. Poisson Analysis of Commuter Flexibility in Changing Routes and Departure Times. *Transportation Research,* Vol. 23B, No. 1, 1989, pp. 53–60.
22. Heckman, J. Sample Selection Bias as a Specification Error. *Econometrica,* Vol. 47, No. 1, 1979.
23. Dubin, J., and D. McFadden. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica,* Vol. 52, 1984.
24. Mannering, F. "Selectivity Bias in Models of Discrete and Continuous Choice: An Empirical Analysis. In *Transportation Research Record 1085,* TRB, National Research Council, Washington, D.C., 1986, pp. 58–62.

Publication of this paper sponsored by Committee on Transportation Supply Analysis.