Exploring Route Choice Behavior Using Geographic Information System-Based Alternative Routes and Hypothetical Travel Time Information Input

MOHAMED A. ABDEL-ATY, RYUICHI KITAMURA, AND PAUL P. JOVANIS

A statistical analysis of commuters' route choice is presented. A binary logit model with normal mixing distribution using stated preference repeated-measurement data is estimated. General descriptive statistics are initially introduced in the paper to explore various route choice criteria and provide the basis for model estimation. The analysis is based on mail-out/mail-back surveys that were customized using routes generated by a geographic information system and information gathered from two previous route choice surveys. The results indicate the significance of travel time, travel time reliability, traffic safety, and roadway characteristics on route choice. The estimation results also underscore the influence of traffic information on route choice.

Fastest-path routing has been adopted over the years because of its simplicity and linkage with algorithms for generating equilibrium in static traffic assignment models. However, in real life, driver's routes are likely to deviate from the fastest path in various ways. Empirical research on route choice behavior shows that drivers use numerous criteria in formulating a route: travel time, number of intersections, traffic safety, traffic lights, and other factors. Drivers' experiences, habits, cognitive limits, and other behavioral considerations may also produce variations in route selection. Viewed in this light, one can see that assuming travel time as the sole criterion of route choice is indeed an unrealistic abstraction of individual driver behavior and when aggregated at the network level may result in an inaccurate representation of traffic.

A number of studies have been performed in the past on route choice. Minimizing travel time is considered the most important criterion affecting drivers' route choice (1–3). Also, directness (2) and less congestion (3) were among the important reasons. Wachs (3) concluded that socioeconomic and demographic characteristics do not clearly relate to attitudes toward route choice criteria, whereas Jou and Mahmassani (4) and Mannering et al. (5) found that socioeconomic characteristics together with the traffic network were important determinants of route changing behavior.

An important factor that has been introduced frequently in the past few years is traffic information and its effect on commuters' behavior in general and on route choice in particular. Insights into drivers' route choice will help us understand the effect of information, which might also be a factor in new network-level traffic models.

This paper uses data collected from a route choice survey. The survey included two major components: a revealed preference (RP) section based on the attributes and perceptions of the respondents' primary (chosen) and geographic information system (GIS)-generated alternate routes and a stated preference (SP) section using repeated discrete choice scenarios. The main objective of the paper is to explore the criteria that influence commuters' route choice and to investigate the effect of advanced traveler information on route choice. The paper presents general descriptive statistics of commuters' route choice. A binary logit route choice model using stated preference data is also presented.

ROUTE CHOICE SURVEY

An ongoing effort for Partners for Advanced Transit and Highways (PATH) at the University of California, Davis, is to investigate the actual route choices of drivers, with the objective of developing refined route choice models that can include the effect of traveler information.

Two computer-aided telephone interviews (CATIs) were conducted in May 1992 and May 1993, respectively. These surveys investigated the actual routes used by commuters, their awareness of alternative routes, their attitudes and perceptions of several commute characteristics, and the traffic information they acquire and its effect on their route switching and choice. Several previous studies by the authors present the design of the CATI surveys and the data analyses (6–8).

A third route choice customized survey was developed targeting the respondents interviewed in the previous two CATI surveys. This survey contains the GIS-based RP route choice questions and the repeated discrete choice SP questions that evaluate the potential effect of advanced traveler information systems (ATIS). The survey was designed to obtain the following information:

- Route attributes considered important by the individual in the decision process that leads to the choice of a route;
- Commuter familiarity with highway and street networks and its potential effect on route choice;
- Commuter willingness to use ATIS; and
- Effect of advanced traffic information on route choice.

Response Rate

The number of targeted respondents was restricted by the availability of their addresses and the success in geocoding their home and
work locations using a GIS. Home and work locations were successfully geocoded, and addresses were available (agreed to provide the address during the second CATI survey) for 263 respondents. The 263 questionnaires were customized according to each respondent's home and work locations, primary route, and travel time. The questionnaire included each respondent's primary route (from the CATI surveys), a GIS-generated minimum path route between the commuter's origin and destination, and SP choice scenarios customized using the commuter's primary route and actual travel time. The questionnaires were sent to the respondents along with a postage-paid return envelope and an incentive of $2.00. A total of 143 respondents completed and returned the questionnaires (54.4 percent response rate).

Survey Design

The revealed preference section and the stated preference section, both heavily customized for the respective respondents, are described.

Revealed Preference Section

The main objective of the revealed preference section is to understand why commuters choose a particular route (in this case their primary—or most frequently used—route); why they do not necessarily use the GIS-generated "optimal" route; how they perceive the primary and optimal routes; how familiar they are with the street and highway network; and how willing they are to use and accept the advice of an ATIS.

The primary commute route for each respondent is identified from the previous CATI surveys. Each segment of the primary route is presented to the respondent in a table; then the respondent is asked to rate a series of subjectively measured route attributes related to the primary route.

On the basis of each respondent's origin (home) and destination (work), and using GIS capabilities, the Navigation Technology's data bases are used to generate optimal routes. Navigation Technology's data bases are detailed data bases that include all the highways and streets in the study area. The optimal route is presented to the respondent in the questionnaire, followed by several questions that measure the respondent's familiarity with this route, willingness to use an ATIS, and rating of a series of route attributes. The RP data provide significant insights into the factors that influence route choice.

Stated Preference Section

The main objective of the SP section is to investigate the effect of ATIS together with roadway type, travel time, and familiarity with a particular route, on the route choice. SP methods become an attractive option in transportation research when revealed preference methods cannot be used in a direct way to evaluate the effect or demand for nonexisting services (e.g., ATIS). SP methods are easier to control, more flexible, and economical as each respondent may provide multiple observations for variations of the explanatory variables.

In this survey, each respondent is provided with three scenarios; in each, the respondent has to choose between two routes (Figure 1 shows an example of one of the scenarios). The choice is binary: Route 1 is customized for each respondent so that the SP design would be as realistic as possible, whereas Route 2 is hypothetical. For Route 1 it is stated: "Your primary route using . . . " and then a segment of the respondent's actual route is written. The travel time of Route 1 is the respondent's actual commute time as stated in the CATI surveys, and the road type is the actual route type of the primary route (mainly freeway, mainly surface streets, or freeway and surface streets). The objective is to use the route that the respondent is familiar with and make the SP design realistic. Road type of Route 2 is one of the following: mainly freeway, mainly surface streets, or a combination of freeways and surface streets.

For the travel time on the alternative route to be as realistic as possible, and because both routes have the same origin and destination, the travel time on both routes is likely to be close to a great extent. Therefore, normal travel time on Route 2 is as follows:

- 0.9 * (normal travel time on Route 1)
- 1.0 * (normal travel time on Route 1)
- 1.1 * (normal travel time on Route 1)

Traffic information is available on either Route 1 or Route 2, but not both. If traffic information is available an estimation of the travel time on that day is one of the following:

- 0.9 * (normal travel time on the same route)
- 1.0 * (normal travel time on the same route)
- 1.1 * (normal travel time on the same route)
- 1.2 * (normal travel time on the same route)
- 1.4 * (normal travel time on the same route)

These values are chosen to be as realistic as possible to represent light and usual traffic conditions (factors of 0.9–1.1), mild traffic conditions (factor of 1.2), and heavy traffic conditions that might be caused because of, for example, an accident (factor of 1.4).

If the information system estimates an above-normal travel time, the cause of the delay is given to the respondent. The cause of the delay is either accident, maintenance, stalled vehicle, or regular congestion. ATIS were defined to the respondents as a system that can offer personalized information about a trip and give advice about other routes to take while considering current traffic conditions.

All possible combinations of the previous cases are considered, after excluding the obvious choices (e.g., if Route 1 is faster and has information that predicts no delays). In all, 68 different combinations were used, three for each respondent randomly.

FACTORS AFFECTING ROUTE CHOICE

As mentioned earlier, one of the main objectives of this study is to determine which route attributes are considered important by the individual in the decision process that leads to the choice of a route. Respondents were asked to rank several factors that made them choose their primary route. The factors and the respondents' rankings are given in Table 1. Shorter travel time is the most important factor (first reason) for choosing the primary route (ranked as the
First reason by 40 percent of the respondents) followed by both
travel time reliability (32 percent) and shorter distance (31 percent).
About 62, 54, and 47 percent indicated that shorter travel time,
travel time reliability, and shorter distance, respectively, as either
the most or second important reason for choosing their primary
route. Other reasons included fewer traffic signals, greater traffic
safety, and lack of unsafe neighborhoods, which about 11, 6, and
4 percent of the respondents, respectively, considered the most
important reason for route choice.

Table 2 indicates the factors that make respondents choose their
primary route over the suggested optimal route (which was gener­
ated using a GIS system). Again, the results support the previous
result that travel time minimization is the most significant factor.
About 62.9 percent of the respondents indicated that they do not use
the suggested optimal route because their primary route is faster.
However, there exist other factors that enter into the decision to
choose a particular route. Shorter distance, travel time reliability,
and traffic safety were among the factors indicated by 37.8, 37.1,
and 28.7 percent, respectively.

Other factors also enter into some individuals’ decision to use a
particular route. Number of roadway segments, freeway use, trip
chaining, neighborhood security, and familiarity were among the
factors less frequently stated. Overall, 10.5 percent of the respon­
dents indicated that the suggested optimal route is the same as their
primary route (they are already using the optimal route).

This result clearly shows that minimizing travel time is the pri­
mary reason for route choice, which conforms to many previous
studies (1–3). This result also illustrates that minimizing travel time
is not the only factor; there exist other important reasons, such as
tavel time reliability. Travel time reliability adds the measure of
uncertainty to the route choice and introduces the significance of an
information system that may help reduce travel time by selecting
routes adaptively. In another paper by the authors in this Record,
travel time variation was found to significantly affect route choice.
Also, this result indicates that shortest-path criteria (either time or
distance) solely are an unrealistic abstraction of individual driver
behavior. It might be more realistic to include all the previous fac­
tors in determining drivers’ route choice behavior and giving each
factor a weight that represents its significance in the route choice.

Figure 2 indicates respondents’ perception of their familiarity
with the GIS-generated route (this measure might indicate the
respondents’ overall familiarity with their streets/highways net­
work). The figure shows that a large majority of the respondents
(73 percent) consider themselves “extremely familiar” with the
suggested route, and 21.6 percent considered themselves “very
familiar” with this route. The rest, about 5 percent, considered
themselves “somewhat familiar” with the route. Only one respon­
dent considered himself “not at all familiar.” Also, about 54.3 per­
cent of the respondents indicated that they had used the GIS-gener­
ated route before and 28.6 percent had used part of the route,
whereas only 17 percent had not used this route.

The previous results indicate that the majority of the respondents
are familiar to a large extent with their networks, which suggests
that the commuters’ unfamiliarity with alternative routes is not one
of the main reasons that they choose a particular route; it is their per­
ceptions of the attributes of a particular route, as discussed earlier
(travel time, travel time reliability, distance, safety, etc.) that lead to
a certain choice.
### TABLE 1 Reasons for Choosing Primary Route

<table>
<thead>
<tr>
<th>Reason for route choice</th>
<th>1st reason</th>
<th>2nd reason</th>
<th>3rd reason</th>
<th>4th reason</th>
<th>5th reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shorter travel time</td>
<td>58 (40.6%)</td>
<td>31 (21.7%)</td>
<td>11 (7.7%)</td>
<td>8 (5.6%)</td>
<td>4 (2.8%)</td>
</tr>
<tr>
<td>Travel time is reliable</td>
<td>46 (32.2%)</td>
<td>31 (21.7%)</td>
<td>21 (14.7%)</td>
<td>14 (9.8%)</td>
<td>8 (5.6%)</td>
</tr>
<tr>
<td>Shorter distance</td>
<td>45 (31.5%)</td>
<td>23 (16.1%)</td>
<td>17 (11.9%)</td>
<td>11 (7.7%)</td>
<td>4 (2.8%)</td>
</tr>
<tr>
<td>Fewer traffic signals</td>
<td>15 (10.5%)</td>
<td>15 (10.5%)</td>
<td>24 (16.8%)</td>
<td>24 (16.8%)</td>
<td>11 (7.7%)</td>
</tr>
<tr>
<td>Greater traffic safety</td>
<td>8 (5.6%)</td>
<td>15 (10.5%)</td>
<td>14 (9.8%)</td>
<td>17 (11.9%)</td>
<td>31 (21.7%)</td>
</tr>
<tr>
<td>No unsafe neighborhoods</td>
<td>5 (3.5%)</td>
<td>7 (4.9%)</td>
<td>9 (6.3%)</td>
<td>11 (7.7%)</td>
<td>19 (13.3%)</td>
</tr>
<tr>
<td>Drive more on carpool lanes</td>
<td>1 (0.7%)</td>
<td>3 (2.1%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Note:** Summing each column might exceed 100%, that is because some people chose two factors as 1st or 2nd reason, e.g., they consider shorter travel time and travel time reliability as the most important reason for route choice.

### TABLE 2 Reasons for Not Using GIS-Generated Optimal Route

<table>
<thead>
<tr>
<th>Reason</th>
<th>No. of respondents (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary route is faster</td>
<td>90 (62.9%)</td>
</tr>
<tr>
<td>Primary route is shorter</td>
<td>54 (37.8%)</td>
</tr>
<tr>
<td>Travel time is unpredictable</td>
<td>53 (37.1%)</td>
</tr>
<tr>
<td>Primary route is safer</td>
<td>41 (28.7%)</td>
</tr>
<tr>
<td>Many short roadway segments</td>
<td>16 (11.2%)</td>
</tr>
<tr>
<td>Primary route involves more freeway segments</td>
<td>14 (9.8%)</td>
</tr>
<tr>
<td>Have to make stop on the way along the primary route</td>
<td>11 (7.7%)</td>
</tr>
<tr>
<td>Primary route does not include insecure neighborhoods</td>
<td>9 (6.3%)</td>
</tr>
<tr>
<td>Not completely familiar with this route</td>
<td>5 (3.5%)</td>
</tr>
<tr>
<td>Had a bad experience in the past with the suggested route</td>
<td>5 (3.5%)</td>
</tr>
</tbody>
</table>

**Note:** Multiple answers are allowed (respondents can choose more than one factor)
ROUTE CHOICE MODEL USING STATED PREFERENCE DATA

Route choice models usually are estimated with observations of actual behavior, or RP data using discrete choice models (as presented in the previous section). However, hypothetical choice scenarios may be needed if the RP data do not provide information on preferences for nonexisting services (e.g., ATIS). In this study the SP scenarios are customized according to each respondent’s route, roadway type, and travel time. This approach made the hypothetical choices realistic to a great extent, which is believed to be useful because a respondent’s choices are more likely to represent the actual behavior.

As mentioned earlier, respondents were presented with three hypothetical scenarios. These scenarios are designed to investigate the respondents’ route choice in the existence of ATIS. The estimation of a logit model with repeated observations for each respondent gives rise to an obvious serial correlation of disturbances. This may be caused by heterogeneity, which refers to variations in unobserved contributing factors across behavioral units. If behavioral differences are largely caused by unobserved factors, and if unobserved factors are invariant over time but correlated with the measured explanatory variables, then estimates of model coefficients will be biased if this heterogeneity is not considered. Even without the correlation between the explanatory variables and unobserved factors, estimates of standard errors will be biased when the disturbances of a series of choices are serially correlated.

Methodological Approach

The approach taken in this paper to account for unobserved heterogeneity is to assume a parametric functional form for the pattern of the heterogeneity. The vector of observed choices or responses for individual \( i \) is defined as \( y_i \). Each element of \( y_i \) is written as \( y_{it} = 1, \ldots, T_i \), each of which is a repeated binary choice, expressed as the integers 0 and 1. The length of \( y_i \) is \( T_i \), which may vary between individuals. The sample size is written as \( I \), so \( i = 1, \ldots, I \).

In the context of the short-term repeated choice sets data analyzed in this paper, it is possible to argue the existence of no state dependence (the utility of one period does not depend on choices of the previous periods) and stationarity (neither the variance of the error term nor the serial correlation between the error terms depend on time) (9,10). The probabilities that individual \( i \) chooses alternatives 0 and 1, \( p_{0i} \) and \( p_{1i} \) respectively, are given as

\[
p_{0i} = P(y_{it} = 0 | \alpha, \beta, x_{it}) = 1/[1 + \exp(x_{it}'\beta + \alpha)]
\]

\[
p_{1i} = P(y_{it} = 1 | \alpha, \beta, x_{it}) = \exp(x_{it}'\beta + \alpha)/[1 + \exp(x_{it}'\beta + \alpha)]
\]

where

\( \alpha = \text{constant} \),

\( \beta = \text{vector of parameters} \), and

\( x_{it} = \text{vector of exogenous variables} \).

The influence of the unobserved variables in Equation 1 is represented by the constant term \( \alpha \); that is, the influence is assumed constant across individuals. The probability of observing \( y_i = (y_{i1}, \ldots, y_{iT}) \) given \( T_i \) in this specification is

\[
P(y_i | \alpha, \beta, T_i, x_{it}) = \prod_{t=1}^{T_i} \left[ \frac{\exp(x_{it}'\beta + \alpha)}{1 + \exp(x_{it}'\beta + \alpha)} \right]^{D_{it}}
\]

where

\[
D_{it} = \begin{cases} 
1 & \text{if } Y_{it} = 1 \\
0 & \text{otherwise} 
\end{cases}
\]

Heterogeneity is introduced into the model by assuming that the probabilities \( p_{0i} \) and \( p_{1i} \) are conditional on both \( x_{it} \) and an individual specific error term, \( \xi_i \), which represents all the other influences. Equation 1 becomes

\[
p_{0i} = P(y_{it} = 0 | \beta, x_{it}, \xi_i) = 1/[1 + \exp(x_{it}'\beta + \alpha + \xi_i)]
\]

\[
p_{1i} = P(y_{it} = 1 | \beta, x_{it}, \xi_i) = \exp(x_{it}'\beta + \alpha + \xi_i)/[1 + \exp(x_{it}'\beta + \alpha + \xi_i)]
\]

The \( \xi_i \), \( i = 1, \ldots, I \) are assumed to be identically distributed with density function \( f(\xi_i) \) independent of the \( x_{it} \), so that Equation 2 becomes
The model shows that commute distance has a significant effect on route choice. The positive coefficient of the log of commute distance on the actual primary route (Route 1) indicates that respondents with longer distances tend to choose the alternative route (Route 2). This indicates that people with long commutes are more disposed to trying out an alternative route in an attempt to minimize their trip. The use of the log transformation indicates that this effect is nonlinear, with marginal increases in distance playing a stronger role in shorter commutes.

Finally, the significance of $\sigma$ illustrates that the unobserved influences affecting a specific individual’s choice are correlated from one of his selections to the next. This demonstrates the need to use a methodology, for example the normal mixing distribution in this study, to account for unobserved heterogeneity.

**CONCLUSIONS**

This paper is based on data collected from a route choice survey. The survey utilized innovative methods in studying route choice behavior by customizing mail-out/mail-back questionnaires. A GIS was used to generate optimal routes to understand drivers’ familiarity with the highways/streets network and to study commuters’ perceptions of this route in a way that helps identify the factors that influence route choice behavior. The survey included also a customized stated preference section that enables the investigation of the possible impact of ATIS on route choice.

The analysis showed clearly that minimizing travel time is the most important reason for choosing a commute route. About 40 percent of the respondents indicated that shorter travel time is their principal reason for choosing their primary route, and 63 percent indicated that they choose their primary route over the suggested optimal route because their primary route is faster.

However, minimizing travel time is not the sole reason for route choice. A large number of the respondents indicated the significance of other factors, such as travel time reliability, which illustrates the significance of the uncertainty measure in route choice and introduces the significance of an information system that reduces the level of uncertainty and helps commuters select routes adaptively. Other important factors that influence route choice are travel distance and the traffic safety on the chosen route.

Other factors appeared to enter into the route choice process, such as the number of traffic signals and stop signs and neighborhood security. These results suggest that route choice selection is a function of several factors, in which travel time would be assigned a heavy weight, and the other factors would contribute to the function according to the degree in which they influence the route choice. The survey results indicated that most respondents were familiar with the GIS-based alternative routes. Commute route choice appears to be a well-informed choice.

Modeling route choice asserted the significance of travel time on route choice, and showed clearly that ATIS has a great potential in influencing commuters’ route choice even when advising a route different from the usual one. Several other commute factors were found to affect route choice, for example number of different roadway segments, freeway use, commute distance, and travel time reliability.

This paper illustrates clearly that several significant factors influence route choice, including advanced traffic information that provides...

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

ACKNOWLEDGMENTS

The authors thank Navigation Technologies for providing the software and data bases used in developing the survey. Special thanks to the California Department of Transportation and PATH for funding this research.

REFERENCES

4. Jou, R., and H. Mahmassani. Comparability and Transferability of Commuter Behavior Characteristics Between Cities: Departure Time and travel time estimates for commuters. However, more work should be done to study the effect of factors that were not included in this study, along with drivers' experiences, habits, cognitive limits, and other behavioral considerations. An important factor that was raised in this paper that needs more investigation is the effect of travel time variation and uncertainty on route choice. Developing route choice models using the RP and SP data jointly also remains as a future task.

Note: model coefficients are defined for route 2