Discrete/Continuous Analysis of Commuters' Route and Departure Time Choices

SAMEER A. ABU-EISHEH AND FRED L. MANNERING

An analysis of commuters' choices of routes and departure times is undertaken using a discrete/continuous econometric modeling structure. The modeling system is estimated with morning work trip data collected in State College, Pennsylvania. The estimation results provide interesting insights into the behavioral aspects motivating route and departure time choices and underscore the need for proper econometric specification in discrete/continuous model structures. Overall, the model estimations provide surprisingly good fits and show promise for applications in a traditional user equilibrium framework.

The problem of peak-period traffic congestion has served as the primary motivation for countless research studies. Such efforts have ranged in focus from highway capacity analyses, including intersection studies and coordinated traffic signal strategies, to studies of trip-making behavior. The congestion remedies suggested by these studies have given rise to new highway construction, various traffic capacity improvements (e.g., highway widening and improvement and automation of signal timings), high-occupancy-vehicle lanes, car- and vanpooling, and marketing efforts supporting public transportation. Although such remedies have met with varying degrees of success, peak-period traffic congestion continues to be one of the most persistent problems facing the transportation profession.

In this paper, an important behavioral aspect of the peak-period congestion problem, travelers' choices of route and departure time, is considered. In the past few years, there has been a healthy level of research devoted to the general area of departure time and route choice. For example, an econometric modeling approach was adopted by Cosslett (1) and Abkowitz (2) for departure time choice, and by Hendrickson and Plank (3) for the choice of departure time and mode. Hendrickson and Kocur (4), Hendrickson et al. (5), De Palma et al. (6), and Mahmansani and Herman (7) studied departure time in the context of user equilibrium for a single route. Extensions of some of these modeling efforts resulted in the inclusion of a route choice component, but on a limited scale as expressed by the equilibrium approaches of De Palma et al. (6), Mahmansani and Herman (7), and by the econometric simulation approach of Ben-Akiva et al. (8). Although a number of past studies have provided valuable insight into the route and departure time choice decision-making process, they have all developed methodologies that treat departure time as a discrete variable as opposed to a continuous one. The argument in support of the discrete treatment of this variable is that travelers can only distinguish among a few prevailing traffic conditions over a specified departure period. However, by discretizing departure time an arbitrary structure of time intervals is being imposed on the decision model. In this paper, a model that treats departure time as a continuous variable and thereby avoids any a priori restrictions on the modeling approach is developed.

ECONOMETRIC FRAMEWORK

In developing an appropriate econometric structure, first a probabilistic route choice model for travelers' automobile commutes to work is specified. Let the utility provided by each route be a linear function for each traveler,

\[ U_i = \beta (ETT_i) + \sigma (RC_i) \]  

where

\[ U_i = \text{utility provided by Route } i \text{ to the traveler}; \]
\[ ETT_i = \text{expected travel time on Route } i; \]
\[ RC_i = \text{vector of route specific characteristics such as number of traffic signals, queue lengths, and so on, for Route } i \text{ and} \]
\[ \beta, \sigma = \text{estimable parameters}. \]

If a disturbance term is added to Equation 1 such that \( V_i = U_i + \epsilon_i \) is assumed to be distributed with a generalized extreme value (GEV) distribution, it can be shown (9) that the route choice probabilities are given by the standard multinomial logit model,

\[ P(i|R) = \frac{\exp U_i}{\sum_j \exp U_j} \]  

where \( P(i|R) \) is the probability of selecting Route \( i \) from the set of available Routes \( R \), and other terms are as previously defined.

With the route choice model specified, the specification of a continuous departure time model can be considered. The departure time, for each commuter, is defined from the following identity:

\[ \text{With the route choice model specified, the specification of a continuous departure time model can be considered. The departure time, for each commuter, is defined from the following identity:} \]
\[ DT = WST - TT - WAT - DC \]  \hspace{1cm} (3)

where

\[ \begin{align*}
DT &= \text{departure time;} \\
WST &= \text{work start time;} \\
TT &= \text{travel time;} \\
WAT &= \text{work access time (i.e., walking time from parking location to work location); and} \\
DC &= \text{delay cushion defined as the time difference between } WST \text{ and arrival time (i.e., } DC = DT + TT + WAT). \\
\end{align*} \]

For the purposes of this analysis, it is assumed that \( WST \) and \( WAT \) are exogenous to the route and departure time choice process. Therefore, the aspects of departure time determination that are controllable by commuters include travel time and the delay cushion. In similar work, Mahmassani and Chang (10) analyzed the individual's departure time through a dynamic boundedly rational framework, but one that was based only on a schedule-delay acceptability mechanism.

Most previous work has viewed route travel time as a factor beyond a commuter's control. In other words, under specified flow conditions, all commuters face the same route travel times. However, in reality, this assertion is valid only under extremely congested conditions. Under most flow conditions, individual commuters have considerable control over their travel times by their abilities to alter driving speeds, risk-taking behavior, and reaction times in the traffic stream and at intersections. Given this condition, a linear model of individual commuters' choice of travel time is defined by

\[ TT_i = \tau + \alpha (RC_i) + \theta (SE) + \eta (VC) + \nu_i \]  \hspace{1cm} (4)

where

\[ \begin{align*}
TT_i &= \text{work trip in-vehicle travel time (min) on Route } i; \\
RC_i &= \text{vector of route specific characteristics for Route } i \text{ (e.g., flow rate);} \\
SE &= \text{vector of commuters' socioeconomic characteristics of the commuter;} \\
VC &= \text{vector of vehicle characteristics used to commute;} \\
\nu_i &= \text{disturbance term; and} \\
\tau, \alpha, \theta, \eta &= \text{estimable parameters.} \\
\end{align*} \]

From an econometric perspective, the estimation of Equation 4 gives rise to a classic problem of selectivity bias because route and travel time choices are interrelated decisions. To illustrate this problem, consider an origin-destination pair connected by two routes, one a freeway and the other an arterial. It is unrealistic to assume that the travel time behavior of people observed to be using the freeway will be identical to that of people using the arterial. Observed freeway users may tend to be faster drivers, in general, because the freeway route offers them the potential to drive at much higher travel speeds. Hence, on the basis of observed route users, a censored sample exists because there is no way to know how fast a freeway user would have driven had he or she selected the arterial or how fast an arterial user would have driven had he or she selected the freeway. Estimation results are biased because users observed on any specific route represent a nonrandom sample formed from a systematic route selection process.

Numerous econometric methods have been developed in recent years to correct such a selectivity bias problem. A comprehensive review of such methods as they apply to transportation is presented by Manncring and Hensher (11). The method selected for use in this study is the expected value method, which has been successfully applied by Dubin and McFadden (12) and by Manncring and Winston (13). To apply this method, Equation 4 is rewritten, conditioning on the choice of Route \( i \),

\[ TT = \tau + \alpha \sum_{k=1}^{n} (RC_i) \phi_{ki} + \theta (SE) + \eta (VC) + \nu \]  \hspace{1cm} (5)

where \( TT \) is the travel time conditional on choice Route \( i \), \( n \) is the total number of route alternatives, and \( \phi_{ki} \) is an indicator variable that equals 1 when \( k = i \) and 0 otherwise.

To arrive at consistent estimates of Equation 5, the choice indicators \( \phi_{ki} \) are replaced by the estimated probabilities from the route choice model (Equation 2). Thus, every route-specific variable included in the travel time equation is replaced by its expected value, which is the summation over all route alternatives of the route's selection probability multiplied by its corresponding route-specific attribute.

The remaining component of Equation 2 necessary for departure time prediction is the delay cushion, which is defined as the difference between the work start time and the actual commuter arrival time. A linear model is defined,

\[ DC_i = \eta + \psi (RC_i) + \lambda (SE) + \delta (PREF) + \omega_i \]  \hspace{1cm} (6)

where

\[ \begin{align*}
DC_i &= \text{delay cushion (min) on Route } i; \\
RC_i &= \text{vector of route specific characteristics,} \\
SE &= \text{vector of commuters' socioeconomic characteristics,} \\
PREF &= \text{commuters' preferences for early or late arrivals,} \\
\omega_i &= \text{a disturbance, and} \\
\eta, \psi, \lambda, \delta &= \text{estimable parameters.} \\
\end{align*} \]

As was the case with the travel time model, the delay cushion model must also be corrected for possible selectivity bias, because it is unreasonable to assume that the selection of route and delay cushion are independent decisions. Consistent estimates of Equation 6 are obtained as discussed for the travel time model by replacing all route-specific independent variables by their expected values (see Equation 5).

**EMPIRICAL SETTING AND SAMPLE DESCRIPTION**

To estimate the specified route, travel time, and delay cushion models, a survey of morning commuters in the State College, Pennsylvania, metropolitan area was undertaken. To simplify data collection and subsequent empirical analysis, one origin-destination pair was evaluated. The origin was a large residential development in suburban State College and the destination...
was the Pennsylvania State University and surrounding downtown State College. The residential area comprises mostly multiple-story apartment complexes and the destination is a highly concentrated area of educational and business activity comprising less than 1 m².

Three distinct and diverse routes connect the selected origin-destination pair. One is a four-lane major arterial with center turning lane and a 35-mph posted speed limit. Another is a two-lane rural highway with 12-ft lanes, 4-ft paved shoulders, and a design speed of 45 mph. The third is a four-lane expressway designed to Interstate standards. This diversity of routes makes the selected origin-destination pair particularly well suited to route and departure time choice modeling.

The morning commute survey was designed as a trip log in which respondents provided a variety of information on their most recent work trip, including route choice; make, year, and model of car used; maximum driving speed; departure time; work arrival time; scheduled work starting time (if any); preferred arrival time at work; safety belt use; automobile occupancy; and walking time from parking lot to work location. In addition, general socioeconomic information was collected, including income, age, occupation, marital status, and number of children. In the last week of April 1986, the survey was administered as a postage-free mailback questionnaire to 505 randomly selected origin residents. Although there was nearly a 40 percent response rate, improperly completed forms and respondents with work start times outside of the studied morning peak period produced 151 usable observations. The summary statistics for the usable sample are presented in Table 1.

Table 1 reflects the relatively short commute times that are typical for such a small metropolitan area. The socioeconomic characteristics are also typical for the graduate students and young professionals that dominate the survey sample.

In addition to the commuter survey, extensive traffic-related data were collected for each of the three routes connecting the origin and destination. This information included flow rates, intersection queue lengths, peak-hour volumes, route lengths, and traffic signal characteristics (phasing, cycle times, etc.). All of this information is potentially useful for the route characteristic variables specified in Equations 1, 4, and 6.

ESTIMATION RESULTS

The route choice model as specified in Equations 1 and 2 was estimated first. The estimation results of this model are presented in Table 2. As mentioned earlier, this model specifies the probability of a commuter's selecting one of the three alternate routes (i.e., arterial, rural route, or expressway) described in the previous section.

The first variable included is the route's expected travel time. Expected travel time is defined to be the travel time predicted by the Bureau of Public Roads' (BPR) equation

\[ ETT = T_0 \left[ 1 + \alpha \left( \frac{V}{C} \right)^{\beta} \right] d \] (7)
where

\[ ETT = \text{expected travel time (min);} \]
\[ T_0 = \text{travel time (min/mi) at zero flow (i.e., at speed limit);} \]
\[ V = \text{peak-hour volume (veh/hr) measured from field surveys;} \]
\[ C = \text{capacity of the route;} \]
\[ d = \text{distance from origin to destination;} \]
\[ \alpha, \beta = \text{route specific parameters that are functions of speed limit and capacity.} \]

The values of \( \alpha \) and \( \beta \) used in estimation are obtained from the Branston study (14) in which transferable values are presented for different highway types.

The expected travel times calculated by Equation 7 may vary from commuter to commuter because their precise destinations within the general Pennsylvania State University and downtown State College destination area are considered. In other words, travel on local access streets is considered in addition to the three line-haul routes discussed earlier. Expected travel times are therefore distance weighted to account for travel on highway segments of varying capacity and speed limit. All variables in the estimated models take into account this precise destination consideration.

As a final point, it is important to note that the use of expected travel times as defined in Equation 7 avoids endogeneity problems that would be encountered if actual travel times were used. This problem arises because travel time and route choice decisions are interrelated and a correlation between an explanatory variable (travel time) and the disturbance term would exist. [See Mannering and Hensher (11) for a discussion of this problem in a discrete/continuous modeling framework.]

Returning to the coefficient estimates presented in Table 2, \( ETT \) has the anticipated negative effect on route selection probabilities. Moreover, estimation results indicate that higher-income commuters (those earning $30,000 or more) find travel time to be more onerous than their lower income counterparts. This presumably reflects the higher value of time of high-income commuters.

The percentage of coordinated traffic signals has a positive influence on the probability of route selection, indicating that commuters value the reduction in the variance of travel time resulting from signal coordination. Finally, the number of traffic signals increases the probability of route selection for individuals with flexible work start time. This suggests a willingness among flexible-time commuters to accept a higher travel time variance and to gamble for the lowest travel time route. (Recall that expected travel times are explicitly considered in the model.) Statistically, the route choice model performs well with relatively low standard errors and a high degree of log-likelihood convergence.

The travel time model is estimated by ordinary least squares and the dependent variable is the in-vehicle work trip travel time in minutes. The estimation results for uncorrected models and models corrected for selectivity bias (i.e., Equations 4 and 5, respectively) are presented in Table 3. The difference between corrected and uncorrected coefficient estimates underscores the importance of correcting for possible selectivity bias.

For specific coefficient estimates, the expected travel time as defined for the route choice model is a strong predictor of actual travel time. This variable is actually capturing a number

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected travel time if income less than $30,000 (in minutes)*</td>
<td>-0.585</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Expected travel time if income $30,000 or more (in minutes)*</td>
<td>-0.753</td>
<td>(0.155)</td>
</tr>
<tr>
<td>Percent of traffic signals coordinated</td>
<td>0.043</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Number of traffic signals if flexible work start time</td>
<td>0.126</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td>Log likelihood at zero convergence</td>
<td>-241.15</td>
<td></td>
</tr>
<tr>
<td>Log likelihood at convergence</td>
<td>-106.45</td>
<td></td>
</tr>
</tbody>
</table>

*See text for precise definition.
TABLE 3 TRAVEL TIME MODEL ESTIMATES, UNCORRECTED AND CORRECTED FOR SELECTIVITY BIAS (Standard Errors in Parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient Estimate</th>
<th>Uncorrected</th>
<th>Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>8.407</td>
<td>5.079</td>
<td></td>
</tr>
<tr>
<td>Expected travel time (in minutes)*</td>
<td>0.189</td>
<td>0.267</td>
<td></td>
</tr>
<tr>
<td>Flow rate (in vehicles per hour, per lane)*</td>
<td>0.00463</td>
<td>0.00977</td>
<td></td>
</tr>
<tr>
<td>Sex (1 if male, 0 if female)</td>
<td>-0.414</td>
<td>-0.470</td>
<td></td>
</tr>
<tr>
<td>Safety belts (1 if used, 0 if not used)</td>
<td>-1.416</td>
<td>-0.873</td>
<td></td>
</tr>
<tr>
<td>Age (1 if 30 years old or less, 0 otherwise)</td>
<td>-0.677</td>
<td>-0.851</td>
<td></td>
</tr>
<tr>
<td>Vehicle vintage (1 if 5 years old or less, 0 otherwise)</td>
<td>-0.738</td>
<td>-0.555</td>
<td></td>
</tr>
</tbody>
</table>

| Number of observations                       | 151                  | 151         |
| R-squared                                    | 0.150                | 0.216       |

*See text for precise definition.

of physical characteristics of the route, including speed limit, capacity, and distance (see Equation 7). The other route-specific variable in the model is the instantaneous flow rate defined as the observed flow rate over the 5-min interval that occurs 5 min after the commuter's reported departure time, transformed into equivalent vehicles per hour per lane. Understandably, this variable is strongly positive, indicating that increasing instantaneous flow rate increases commuters' travel times.

Three socioeconomic variables were included in the model. The sex indicator variable confirmed a priori expectations that males tended to drive faster than females. The safety belt indicator variable produced a negative coefficient, suggesting that safety belt users tended to drive faster. This result lends some support to Sam Peltzman's famous hypothesis concerning the tendency of safety belt users to drive more recklessly, thereby mitigating the potential benefits of safety belt usage (15). Finally, the age indicator coefficient indicated that younger commuters tend to drive faster than older ones.

The only vehicle characteristic included in the model was the vintage. This variable indicates that newer vehicles, with more sound bodies and steering, are driven faster than older vehicles. Unfortunately, the sample was not large enough to explore the differences among high-performance, compact, intermediate, and large cars.

The delay cushion model was defined only for those respondents that actually had fixed work start times. Because the sample included many researchers and self-employed professionals, only 90 of the 151 respondents had fixed work start times. Because the route choice behavior of these respondents may differ somewhat from the full sample estimates presented in Table 2, the route choice model was reestimated using only the 90 commuters with fixed work start times. The specification is identical to the earlier route choice model, of course excluding the flexible work start time traffic signal variable, and the estimation results are presented in Table 4. The route choice probabilities estimated from this model will be used to produce the expected values needed to correct for selectivity bias in the delay cushion model. For the 61 respondents with no fixed work start times, Equation 3 does not apply. For these commuters, departure times are simply their stated preferred arrival times minus their travel and work access times. The stated preferred arrival time is exogenous to the route and departure
TABLE 4 ROUTE CHOICE COEFFICIENT ESTIMATES FOR PARTICIPANTS WITH FIXED WORK START TIMES (Standard Errors in Parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected travel time if income less than $30,000 (in minutes)*</td>
<td>-0.368</td>
<td>0.073</td>
</tr>
<tr>
<td>Expected travel time if income $30,000 or more (in minutes)*</td>
<td>-0.340</td>
<td>0.157</td>
</tr>
<tr>
<td>Percent of traffic signals coordinated</td>
<td>0.0073</td>
<td>0.011</td>
</tr>
<tr>
<td>Number of observations</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Log likelihood at zero</td>
<td>-98.88</td>
<td></td>
</tr>
<tr>
<td>Log likelihood at convergence</td>
<td>-72.88</td>
<td></td>
</tr>
</tbody>
</table>

*See text for precise definition.

FIGURE 1 The distribution of actual delay cushions.
time choice decision-making process. A similar treatment was used by Mahmassani and Chang (16), for whom preferred arrival times were supplied by commuters.

The dependent variable in the delay cushion model is the difference between the work start time and actual arrival time in minutes and will be positive if the commuter arrives early and negative if the commuter arrives late. The actual observed distribution of these delay cushions is shown in Figure 1. The coefficient estimates for the regression model, both corrected and uncorrected for selectivity bias, are presented in Table 5.

The only route specific variable included in the model is the expected travel time, defined as before. The coefficient is negative, indicating that the longer the commute the less the delay cushion. This negative sign may be an outgrowth of the rather short commuting distances that the sample of travelers experienced. That is, because absolute variance in travel time is fairly small due to the short travel distances, commuters tend to decrease their delay cushions to compensate for longer in-vehicle travel times, knowing that the likelihood of a late arrival is rather small. It would be interesting to reestimate this model with a longer, higher-variance commute and to reassess the delay cushion and expected travel time relationships.

The socioeconomic variables include income and age. The income coefficient is negative, indicating that higher-income people prefer shorter delay cushions (i.e., have a higher value of time). The age coefficient is positive, suggesting that older workers tend to be more risk-adverse by choosing longer delay cushions.

The final variables are commuters’ preferences for early, on-time, or late arrivals. These preferences are accounted for by indicator variables with on-time implicitity set to zero. The coefficients of the preferred-early and preferred-late arrival indicator variables are of plausible sign and are highly significant statistically. Overall, the $R^2$ value of this model is surprisingly high considering the high variance likely to be present in all delay cushion data.

### ESTIMATION NOTES

The inclusion of flow-dependent variables in all models gives rise to two important estimation concerns. The first is one of selectivity in that traveler behavior in response to prevailing flows is observed only for the departure time actually chosen. However, unlike the route choice selectivity bias problem discussed earlier, there is no theoretical basis for assuming that the behavioral characteristics of individuals departing at different times will systematically differ. Thus, although the sample is censored in that individuals are not observed departing at all

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient Estimate Uncorrected</th>
<th>Coefficient Estimate Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.288 (3.979)</td>
<td>0.497 (4.108)</td>
</tr>
<tr>
<td>Expected travel time (in minutes)*</td>
<td>-0.137 (0.161)</td>
<td>-0.268 (0.187)</td>
</tr>
<tr>
<td>Income (in thousands of dollars)</td>
<td>-0.091 (0.082)</td>
<td>-0.091 (0.081)</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>0.174 (0.093)</td>
<td>0.172 (0.093)</td>
</tr>
<tr>
<td>Preferred early cushion (1 if prefer to</td>
<td>9.078 (1.424)</td>
<td>9.119 (1.417)</td>
</tr>
<tr>
<td>arrive before work start time, 0 otherwise)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preferred late cushion (1 if prefer to</td>
<td>-7.047 (3.95)</td>
<td>-7.248 (3.919)</td>
</tr>
<tr>
<td>arrive after work start time, 0 otherwise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>$R$-squared</td>
<td>0.406</td>
<td>0.415</td>
</tr>
</tbody>
</table>

*See text for precise definition.*
available times, there is no reason to believe that it is censored nonrandomly, and as a result selectivity bias and the associated endogenicity of flows are not an issue with respect to departure times.

The second point is one of endogenicity in the more traditional sense. As specified by the equilibrium condition in Equation 1, flow determines demand but is itself an outcome of demand. This relationship can be safely ignored during the estimation of the specified individual choice models because the impact of a single individual's choice on total traffic flow is negligible. Therefore, each individual will view traffic flow as exogenous to the route and departure time choice process.

MODEL APPLICATIONS

The greatest potential application of the route/departure time choice modeling system is in the context of user equilibrium traffic assignment. The modeling system offers the potential to evaluate the traffic-related impacts of a wide range of policy options related to physical changes in the highway system. In addition, due to the behavioral nature of the models, the impacts of shifts in population demographics can also be assessed.

Unfortunately the model structure does not readily lend itself to equilibrium solutions. All three models (route, travel time, and delay cushion) include vehicle volumes as explanatory variables either in expected travel time terms or directly. Although volumes can be considered exogenous when estimating individual choice models as previously discussed, they must be considered endogenous in the context of a user equilibrium. This endogenicity produces a more complex problem than standard route choice equilibriums, such as stochastic user equilibrium (17), due to the presence of travel time and delay cushion models. Equilibrium with such a group of models is technically possible (18) but is beyond the scope of this paper.

SUMMARY AND CONCLUSIONS

Most previous work on route and departure time choice has viewed departure time as a discrete variable. In this paper a route and departure time choice modeling system is developed that treats departure time as a continuous variable. The models were estimated with a sample of work trip commuters and the resulting coefficient estimates were of plausible sign and reasonable statistical significance.

The findings of this study give rise to a number of important points. First, the prospects for continuing advances in development of behavioral route and departure time choice models are most promising. Such models offer the potential for significantly expanding understanding of this critical decision-making process. The second point relates to the importance of proper econometric specification in the estimation of discrete/continuous route and departure time choice models. The estimation results (Tables 3 and 5) suggest that the potential for selectivity bias is considerable. Finally, from the perspective of applications, it is important that future work be directed towards incorporating route and departure time choice models such as those estimated in this paper into a user equilibrium framework. Only then can the true value of such a modeling approach be realized.

REFERENCES