Highway Assignment Method Based on Behavioral Models of Car Drivers' Route Choice

ROBERT W. ANTONISSE, ANDREW J. DALY, AND MOSHE BEN-AKIVA

This paper proposes a highway traffic assignment technique based on an improved behavioral model of drivers' route choice as developed in a recently completed study in the Netherlands. Route choice models are developed from data collected in three corridors in the Netherlands. The models presented here, which are based on evidence of drivers' varying valuations of a number of road characteristics, are (a) probabilistic and (b) based on more variables than were used in previous models. They contrast with the underlying route choice models of conventional traffic assignment procedures, which are typically based on a single measure of travel impedance (e.g., travel time, generalized travel cost). A key feature of the models developed in the present study is that they are based on data describing the actual routes chosen by individual drivers. The paper describes how these models are used to generalize assignment methods through the exploitation of a multiclass-user technique. In an uncongested network, several routes typically will be predicted to be used between a given origin and destination; as congestion increases, so will the diversity of routes used. Several models appropriate for use in varying circumstances of data availability are presented and compared. Model inputs (e.g., road attribute data) are described, and practical implications of the underlying structural assumptions are discussed. Spatial transferability of the models is appraised on the basis of the differing results obtained for the three corridors studied. Finally, advantages and limitations of application of the proposed assignment method compared with conventional procedures are discussed.

A central element of the traffic assignment procedure is a model of the traveler's decision about which route to take given the origin, destination, and mode of travel of a trip. The problem of route choice for a traveler might be stated as follows: Given the other characteristics of the trip to be made—purpose, time, origin, destination, and mode, for instance—choose the "best" route through the transportation network in terms of some criterion. This best route is most often thought of as the one that minimizes travel disutility. Existing traffic assignment models often assume single measures of travel disutility such as travel time or distance, or some simple formula of generalized travel cost.

R. W. Antonisse, Operations Directorate, Massachusetts Bay Transportation Authority, Room 4730, Ten Park Plaza, Boston, Mass. 02116. A. J. Daly, Hague Consulting Group, b.v., Surinamestraat 4, 2585 GJ The Hague, Netherlands. M. Ben-Akiva, Department of Civil Engineering, Massachusetts Institute of Technology, Room 1-181, 77 Massachusetts Avenue, Cambridge, Mass. 02139.

In reality, the problem of route choice faced by an automobile driver is very complex because of

1. The large number of possible alternative routes through even modestly sized road networks, and
2. The complex patterns of overlap between the various route alternatives.

Realistic replication of the human decision process in route choice—which synthesizes many factors about the trip and the various possible routes in making a choice—with a mathematical model is difficult at present because of the limited understanding of the route choice phenomenon, as well as limited techniques and computational resources.

The primary interest of studying car drivers' route choice is in improving traffic assignment procedures. In particular, accurate predictions of the usage of proposed new infrastructure are essential to the evaluation of the need for that infrastructure. The results of the route choice study suggest that current methods may underestimate the traffic attracted to major new roads. Secondarily, understanding route choice is valuable in attempting to redirect traffic streams so as to make the best possible use of existing roads. The current study is one of very few directed to a better fundamental understanding of this important aspect of behavior and the implementation of that understanding in practical planning methods. The overall objectives were twofold:

1. To improve current understanding of drivers' route choice preferences, and
2. To develop a practical traffic assignment model that reflects this choice process with greater sophistication.

FACTORS AFFECTING ROUTE CHOICE BEHAVIOR

A major task completed during the first phase of this project was an extensive literature review of factors affecting drivers' route choice preferences. Ben-Akiva et al. (*1*) synthesized the results of this review as a set of hypotheses that may be broken down into the following three categories:

1. Drivers' knowledge about alternative routes: Several authors hypothesize that drivers plan their trips in a hierar-

chical fashion, building up from lowest-level (local) roads near the origin of the trip to expressways at the highest level, which they use for the bulk of travel, and back to local streets at the end of the trip (*2–4*). Knowledge may be lacking of local road alternatives to expressway portions of trips (*5–7*). Drivers often are unable to evaluate simple characteristics of paths and thus are unable to find the quickest or shortest route (*2, 8, 9*).

2. Decision processes: Various hypotheses assert that drivers either plot out their entire route before departure or make decisions at road junctions as they encounter them independently from previous decisions (that is, they follow a Markov process), or else they use some combination of these two approaches (*10*).

3. Route attributes and preferences: Specific attributes of routes to which drivers are attracted include travel time (*11–14*), distance (*14*), number of traffic signals (*5*), scenery (especially for nonobligatory trips, such as social or recreational ones) (*6,15*), time or distance on limited-access highways (*15*), safety (*11,15*), commercial development, congestion (*15,16*), road quality, and road signing (*17*).

Most of these hypotheses are not reflected in existing traffic assignment models.

NEW MODEL OF ROUTE CHOICE BEHAVIOR

The earlier work on this project documented by Ben-Akiva et al. (*1*) also included the conceptualization of a two-step model of route choice that (a) narrows down the large number of possible route alternatives to a choice set of a few alternatives and (b) chooses a route from this choice set based on the characteristics of the trip, driver, and attributes of the available alternatives. Survey data were collected for a sample of drivers observed to travel between the cities of Utrecht and Amersfoort, including information on the driver and on the trip itself (including the route actually chosen on the survey day). A network model of the corridor was used to generate sets of alternative routes for the sampled drivers, and a large number of route choice models was tested.

The empirical evidence of the first phase of this study showed that factors other than time and distance play a significant role in interurban route choice. For example, several road attributes that one normally associates with major highways—large capacity, restricted access, high hierarchical level, and high speed limit—were found to positively attract route choice. Traffic signals, on the other hand, were found to have a negative effect.

The estimation results demonstrated the feasibility of the two-stage approach to modeling route choice and produced a model that reflects the hypothesized structure underlying route choice behavior. Finally, a number of market segmentation tests demonstrated that trip purpose, frequency, and length can have important influences on route choice.

OBJECTIVES

The results of the second phase of this project are presented. They are based on a new data collection effort that began in 1980 in two other road corridors in the Netherlands. The primary objectives of the second phase were

1. To test the transferability of both stages of the modeling process as developed in the first phase (the method used to generate the set of alternative routes and the choice model) to the other corridors, and
2. To simplify the choice model as a way to enhance the applicability of the model in a wider geographical area and under conditions of limited road network data.

Drivers were surveyed in two different corridors in the Netherlands in the spring of 1980—one between Amsterdam and Purmerend and the other between Arnhem and Apeldoorn. All the corridors offer a number of viable route alternatives for the many trips between the two cities defining each study area. In each case, a cordon of roadside survey points was laid out across the corridor. At some survey points, return-mail questionnaires were handed out to drivers, whereas at other points license plate numbers were recorded and registered owners of the vehicles were sent a return-mail survey form at home. The surveys asked respondents to trace the route they took on a map provided for the day of the sighting. The questionnaires also asked a range of questions about trip and personal characteristics: purpose at origin and destination, frequency of this trip, age, profession, and so on. Meanwhile, network data were collected from engineering sources.

ROUTE CHOICE MODEL FOR TRAFFIC ASSIGNMENT

This section describes the basic methodological requirements and data and computer needs for forecasting route choice behavior using the new approach.

Methodology

Travel behavior in general and route choice behavior in particular can be considered as choosing between discrete, mutually exclusive alternatives. Discrete choice analysis attaches expressions of attractiveness or utility to each of the available choice options. The utility expression of each alternative generally incorporates information on the attributes that may either add to or detract from its attractiveness. It is then assumed that the decision maker will choose the alternative that is most attractive.

With the primary problem in this case being highway route choice, the two major steps in determining behavior are

1. Identifying a set of route alternatives that the driver can choose among, and
2. Making the choice from this set on the basis of the type of driver and trip conditions and the various attributes of the route alternatives.

Because it would be prohibitively time-consuming and behaviorally unrealistic to evaluate the attractiveness of all possible routes between the origin and destination, a method is applied to narrow down the vast number of route possibilities to a few alternatives that may be considered in greater detail.

Once a set of options has been identified, it is necessary to measure the relevant attributes of those options that affect their attractiveness. A choice model is used to relate the probability of choosing each available alternative to its attractiveness, which, in turn, is based on the attributes of the alternatives. For the predictive tool to be successful in forecasting travel behavior under a wide range of circumstances, the choice model must be responsive to how changing travel conditions and varying perceptions affect the relative attractiveness of the available travel options.

Generation of Route Alternatives

As discussed above, the first stage of the route choice modeling process involves the generation of a set of candidate route alternatives from the myriad of feasible paths through the road network. The technique developed in this study is called the "labeling" approach because descriptive labels are attached to the selected route candidates. Each of these labeled routes is optimal with respect to some criterion from among all possible routes between the given origin-destination pair. For example, "quickest," "shortest," and "most scenic" might be criteria used to define three candidates from all route possibilities. The criteria to be used may be extracted from hypotheses regarding influences on route choice behavior and could be considered to constitute a model of drivers' perceptions of a road network.

So that these labels can help determine specific paths through the given network, a quantitative descriptor based on available network data must be selected to measure a route in terms of the label criterion. Labeled paths are defined by an impedance function that depends on one or more link attributes. A separate function is specified for each label criterion to be used. Determining the labeled path for a particular criterion is then simply a matter of calculating the associated impedance for all links in the network and executing a minimum-path algorithm that can efficiently generate labeled paths for a large set of origin-destination pairs. Observed chosen route data are required in selecting the most reasonable set of labels to apply in forecasting route choice. The selected set of impedance functions maximizes the frequency of observed routes included in the set of the corresponding labeled paths.

At this point, it is useful to describe the network data available for this study. Two types of data were used in this analysis: a basic network data base system and sets of extra, detailed link attributes. The Dutch Ministry of Transport maintains a computerized "Basisnetwerk" system consisting of many node and link records that represent the national highway network. This system is used extensively in the Ministry's planning and management functions. Node records include the junction's geographic location. Link records include A- and B-nodes, distance, speed code, and road hierarchy level as attributes. These basic attributes supply sufficient information for the generation of a few important labels.

A large number of detailed road link attributes was gathered for the detailed study area within each of the data collection corridors. Example attributes include road surface type, width of roadway, number of lanes in each direction, zoning type of adjacent land, and presence of various types of facilities along the roadside. A large number of alternative labeled paths could then be generated for any driver traveling either entirely or partially through the detailed study area.

In the first phase of the project, 10 labels were selected for application. These same labels were designated for application to the two new study areas in the second phase. The labels chosen and associated quantitative descriptors are described briefly as follows:

- Minimize time: travel time is calculated from information on the distance and average speed of the link.
- Minimize distance: the distance from the link records is applied.
- Maximize travel on scenic roads: the measure of impedance for the route is time spent driving on roads adjacent to nonscenic land uses—city center, dense residential, or industrial—as determined from percentage of link distance through these types of land use, which is available from detailed attributes.
- Minimize number of traffic signals: for each link, the number of traffic signals was calculated using detailed attribute information and the following formula:

$$\text{No. signals} = \text{no. signals along link} + 0.5\,(\text{total signals at nodes})$$

- Minimize travel on congested roads: detailed attributes allowed calculation of volume:capacity (V/C) ratios for road links in the Phase 1 study area. The descriptor is time spent on roads with high V/C ratios. Unfortunately, link volume data were not available for either the Arnhem-Apeldoorn or the Amsterdam-Purmerend study area, and this label had to be dropped from the analysis of joint data.
- Maximize use of expressways: links were classified as expressways if the network speed code was the maximum, that is, 100 km/hr (approximately 60 mph). Time spent on nonexpressway roads was used as a measure of link impedance in this case.
- Maximize travel on high-capacity roads: the impedance measure is time spent on low-capacity roads, that is, roads that either are less than 9 m (approximately 30 ft) wide or have less than two lanes in either direction.
- Maximize travel in commercial areas: again using land use data from the detailed attributes, time spent in noncommercial areas was calculated on the basis of the distance traveled in any land use area other than cities or industrial areas.
- Maximize road quality: for every link in the study area, a road quality rating is available on a scale of 1 (best quality) to 3 (worst quality). Time spent on poor-quality roads—those with a rating of 2 or 3—was measured.
- Hierarchical travel: each link includes an attribute for road hierarchy level. This label favors travel on the highest-level roads—generally limited-access highways. Two impedance measures were used: (a) time spent on roads of the lowest hierarchy (local roads) and (b) time on roads of moderate hierarchy (main roads of regional importance).

With the label descriptors determined, the next step is the specification of the impedance functions. In the case of the "minimize time" and "minimize distance" labels, this is simply the measure itself. For the other labels, however, it was possible for the optimal route of the criterion to deviate unreasonably from the minimum time path. To mitigate this prob-

TABLE 1 LABEL IMPEDANCE FUNCTIONS AND FINAL COEFFICIENT VALUES

Label Criterion	Link Impedance Function (to be minimized)	Initial Coefficients Utrecht-Amersfoort	Initial Coefficients Amst-Purm & Arnhem-Apel
Min. Time	TIME		
Min. Distance	DISTANCE		
Max. Scenic	TIME + β_1(NON-SCENIC TIME)	2.0	2.0
Min. Signals	TIME + β_2(# SIGNALS)	30 sec.	5 min.[1]
Min. Congestion	TIME + β_3(HIGH V/C TIME)	3.0	not used
Max. Expressways	TIME + β_4(NON-EXP. TIME)	3.0	3.0
Max. Capacity	TIME + β_5(LOW-CAP. TIME)	1.5	2.0
Max. Commercial	TIME + β_6(LOW-COMM. TIME)	1.5	1.5
Max. Road Quality	TIME + β_7(LOW-QUAL. TIME)	2.0	2.0
Hierarchical Travel	TIME + β_{81}(HIER. 1 TIME) + β_{82}(HIER. 2 TIME)	5.0, 100	5.0, 100

[1] The value of 5 min. was not tested for Utrecht-Amersfoort. However increased values (above 30 sec.) did not show a large loss of coverage on that data.

lem, the impedance functions were specified as a weighted sum of the primary criterion measure (e.g., scenic time) and total travel time. The only remaining task is the assignment of relative weights to the two component measures. Table 1 shows the impedance functions, specified with weighting coefficients, as applied in this study.

The label parameters were optimized by finding the set of parameter values that maximized the number of observed chosen routes for the area under study that are matched or "covered" by the label set. A straightforward computer algorithm is used to compare each chosen route with the set of labels developed for that origin-destination pair and determine the existence of a match. Labels were introduced into the label set one at a time, optimizing the label's coefficient(s). A final sensitivity analysis ensured that changing any one parameter did not reduce the total matching score of the set of labels. The final Utrecht-Amersfoort label coefficients are shown in Table 1. A similar analysis was carried out in the other two study areas, using the Utrecht-Amersfoort coefficients as initial values. These values are also shown in Table 1, indicating the minimal changes between the areas in this respect.

Conditional Route Choice Models

After the decision maker's set of alternative routes has been generated, the next step is specification of the utility functions for each of the alternatives. Utility functions include a systematic component—an expression of how independent variables affect the attractiveness of the alternatives scaled up by their respective estimatable parameters—and a random term that accounts for the variability in choice behavior independent of the options' attributes. The systematic utility expressions are usually specified as linear combinations of the independent variables.

The "maximum likelihood" method is used to estimate the values of the utility function parameters. Simply put, this method determines the values of all parameters for which the observed choices are most likely to have occurred. The two most commonly applied probabilistic discrete choice models, logit and probit, differ in their assumptions about how the random term of the utility function is distributed [see, for example, Ben-Akiva and Lerman (*18*)].

The different mathematical properties of the random variables mean that each method has its advantages and disadvantages. Logit-form models are generally more flexible in the feasible number and structure of alternatives in the choice problem, and their parameters can be estimated with considerably less computational burden than those of probit-form models. Logit-form estimation programs are widely available. The package used in the route choice study is ALOGIT (*19*).

Two types of independent variables are considered for inclusion in the utility functions: (a) "level-of-service" attributes for each route including, for example, measures of travel time, distance, and travel cost; and (b) dummy variables that take a value of either 1 or 0 depending on whether an alternative meets certain conditions. In this analysis, dummy variables are used to indicate whether the route corresponds exactly to one or more of the labels considered.

Although these variables are objective measures of the routes themselves, different drivers perceive these attributes differently. The most common bases for these differences in perception may be the characteristics of the drivers themselves (e.g., age or profession) and characteristics of the trip being made (e.g., its purpose and the frequency with which it is made). An attempt is made to capture these differences in perceptions through estimation of models for various segments in the population and examination of the variance in the respective model parameter estimates.

Traffic Assignment

The methods outlined in the preceding sections can be applied by an adaptation of a "multiclass-user" (MCU) procedure. In a standard MCU method, classes are defined a priori as using paths that are minimal with respect to a class-specific impedance function. In the models described in this paper, the assignment procedures define the classes as users of each of the labeled routes. Because the usage of these routes is not known a priori and is dependent on the features of the routes, additional steps have to be introduced into the assignment procedure to apply the model. The procedure advocated is outlined in Figure 1.

An important feature of the procedure outlined is the integration of the new choice modeling approach developed in this study with the "capacity-restraint" methods that have been the subject of many previous studies. This integration means that previously developed algorithms, techniques, and so on, can be retained and current methods can be seen as independent improvement that loses none of the previous gains.

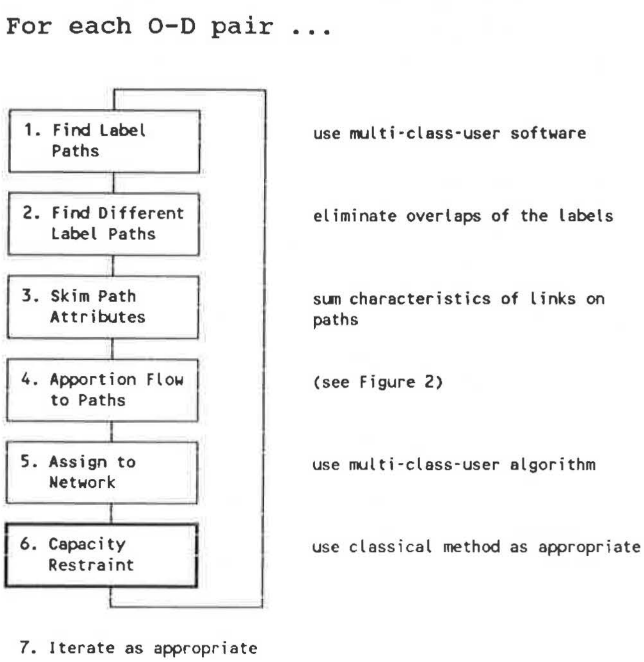

FIGURE 1 Assignment procedure (overview).

The procedure involves six steps for each origin-destination pair for which a positive traffic flow is predicted.

Note that the first, third, fifth, and sixth steps, which are the most demanding in terms of computer processing, are standard MCU assignment steps and are already provided in standard packages. The second step is a simple programming task.

The fourth step in the process shown in Figure 1 is novel and is illustrated in greater detail in Figure 2. For each origin-destination pair, an apportionment is made by the model to each of the labels.

Figure 2 provides for a matrix of size (labels * segments) to be calculated for each origin-destination pair. It may be helpful to note that this procedure would be equivalent to a simple MCU procedure if labels and segments were identified, that is, if the matrix was simply 1.0 on the diagonal and zero elsewhere. The computation necessary to calculate and apply the matrix is not excessive.

In summary, an assignment procedure is proposed that requires comparatively minor extensions to existing software. Execution of this procedure requires little more computer time than a standard MCU method. The procedure is organized as a generalization of existing capacity restraint procedures, thus offering an advance without eliminating the possibilities resulting from previous studies.

MODELING RESULTS

This section summarizes the major quantitative findings from this project. The first subsection discusses results from the choice set generation using the labeling approach described above. The following subsections report and evaluate the final choice models that consider sets of six or fewer route alternatives.

Label Set Coverage of Observed Chosen Routes

With the primary objective of development of an assignment tool that can be applied in all three areas studied in this project, the parameters of the full set of nine labels are developed by maximizing matches of the chosen routes in all three study areas. More manageable six-label and four-label sets were developed for use in the choice modeling stage of the analysis. The labels included in these reduced sets were selected in part on the basis of the expected availability of the link attribute data required for generation of the label.

A computer network analysis package, SATURN (*20*), was used to build the labels between all chosen origin-destination pairs in the three study area networks. A separate computer program was written to compare each observed chosen route with the set of corresponding labeled routes and to summarize the match results. The label parameters developed in the first phase of this project (see Table 1) were used as initial values. Parameters were adjusted one by one, keeping the others fixed, in the direction that increased the number of matches to chosen routes.

Table 2 shows the match results for the initial full-label set and the six-label set using both initial and final label coefficient values. Each row in the main body of the table refers to one label and reports the coefficient value(s) and the set

For a given O-D pair ...

$$
\begin{array}{l|ccccc|l}
\text{Volumes:} & N_1 & N_2 \;\cdots & N_i & \cdots & N_s & \\
 & \text{Seg.1} & \text{Seg.2} \;\cdots & \text{Seg.i} & \cdots & \text{Seg.s} & \text{Flow on paths:} \\
\hline
\text{Path 1} & p_{11} & p_{21} \;\cdots & p_{i1} & \cdots & p_{s1} & V_1 = \Sigma^s_{i=1} \; N_i \cdot p_{i1} \\
\text{Path 2} & p_{12} & p_{22} \;\cdots & p_{i2} & \cdots & p_{s2} & V_2 = \Sigma^s_{i=1} \; N_i \cdot p_{i2} \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
\text{Path j} & p_{1j} & p_{2j} \;\cdots & p_{ij} & \cdots & p_{sj} & V_j = \Sigma^s_{i=1} \; N_i \cdot p_{ij} \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
\text{Path r} & p_{1r} & p_{2r} \;\cdots & p_{ir} & \cdots & p_{sr} & V_r = \Sigma^s_{i=1} \; N_i \cdot p_{ir} \\
\hline
\end{array}
$$

Where: N_i is the number of vehicles for this O-D for segment i (input data);

p_{ij} is the probability of choosing path j for segment i (derived from model and path attributes);

V_j is the predicted volume for path j for this O-D (output to assignment stage).

FIGURE 2 Path apportionment.

of match scores for that single label in the various study areas—Utrecht-Amersfoort (full-label set only), Arnhem-Apeldoorn, and Amsterdam-Purmerend. Three types of match scores are reported for each label in the table:

1. Absolute matches: the total number of chosen paths matched by this individual label for corresponding origin-destination pairs.
2. Incremental matches: the percentage of chosen routes not matched by previous labels but matched by this label for corresponding origin-destination pairs.
3. Marginal matches: the percentage of chosen routes matched by this label and not matched by any other label in the table.

It is clear from Table 2 that a significant gain in coverage of the observed route choices can be realized by including additional criteria besides "minimize time." A comparison of the coverage in percentage terms of the time label alone versus the final six-label and full-label sets yields the results for the three study areas shown in Table 3, in which a four-label set—comprising time, distance, signals, and hierarchy labels—is also presented. Table 3 also shows the decreasing "rate of return" from increasing the size of the label set. Note that the apparent lower coverage of the full-label set relative to the six-label set for the Amsterdam-Purmerend area is explained by the use of the initial, nonoptimal set of coefficient values.

A sensitivity analysis of the label coefficients near their initial values showed that the matching rates generally remained stable. Nevertheless, some gains were made possible for the Arnhem-Apeldoorn and Amsterdam-Purmerend study areas by adjusting the parameters for the "minimize signals" and "maximize capacity" labels. These adjustments are reflected for the six-label set in Tables 2 and 3 and account for the apparent decrease in coverage shown for Amsterdam-Purmerend in the latter table when progressing from the six- to the nine-label set. The match score results in Table 2 as well as data availability considerations were used to decide which labels were to be eliminated to form the reduced sets.

The analysis found that many aspects of the labeling methodology were transferable between the three areas studied. The values of the label parameters, when optimized on chosen routes for the three study areas, also agreed very closely for most of the labels, as can be seen from Table 1.

Choice Modeling Results

Extensive discrete choice modeling was conducted on sets of six and fewer labels. Most of the modeling was done on the combined set of chosen route data from the Arnhem-Apeldoorn and Amsterdam-Purmerend study areas. Alternative specifications tested the explanatory power of various combinations of level-of-service variables as well as various forms for the constants in the utility functions of the alternatives.

A number of model runs explored the effects of applying separate models for various subgroups in the population. Information from the survey responses was used to assign individual drivers to categories of trip length, trip frequency, and trip purpose. A surprising result from the choice modeling analysis was the relatively significant effect of geographical area that could not be explained in terms of differences in trip purpose, length, or frequency profile for the study areas.

Tables 4 and 5 show, respectively, the six-label and four-

TABLE 2 NUMBERS OF CHOSEN ROUTES MATCHED BY SIX AND NINE LABELS

Label	β Coef. Value(s)	Match Scores: Utrecht-Amersfoort Absolute	Utrecht-Amersfoort Incremental %	Arnhem-Apeldoorn Absolute	Arnhem-Apeldoorn Incremental %	Arnhem-Apeldoorn Marginal %	Amsterdam-Purmerend Absolute	Amsterdam-Purmerend Incremental %	Amsterdam-Purmerend Marginal %
Labels not requiring β parameters									
Time		1505	69.9	1659	56.1	0.3	1236	67.4	1.4
Distance		462	1.6	1905	13.6	12.1	357	0.5	0.5
Labels requiring β parameters: initial values									
Scenic	2.0	770	5.7	1288	1.3	1.1	115	0.3	0.2
Signals	30 sec.	851	2.7	1639	0.1	0.1	1233	0.1	0
Capacity	1.5	1058	2.5	1497	0.3	0.3	1170	3.1	0.4
Hierarchy	5.0, 100	712	3.3	1299	0.5	0.5	1246	6.1	6.1
Total 6 labels		1846	85.8	2179	72.0		1436	77.5	
Labels not included in models									
Quality	2.0	1677	0.4	1435	1.6	1.6	1237	0	0
Commercial	1.5	1506	0	1664	0.1	0.1	887	0.8	0.8
Expressway	3.0	501	0.1	1499	0.0	0.0	1197	0	0
Total 9 labels		1857	86.3	2179	73.7		1436	78.3	
Labels requiring β parameters: final values									
Scenic	2.0			1288	1.3	1.1	115	0.3	0.2
Signals	5 min.			1639	0.3	0.3	1233	4.5	3.4
Capacity	2.0			1497	1.8	1.8	1170	3.4	0.8
Hierarchy	5.0, 100			1299	0.5	0.5	1246	5.1	5.1
Total 6 labels				2179	73.6		1436	81.2	

TABLE 3 COMPARISON OF LABEL SET COVERAGE

Study Area	Time Only (%)	Four Labels[a] (%)	Six Labels (%)	Full-Label Set[b] (%)
Utrecht-Amersfoort	69.9	N/A[c]	85.8	86.3
Arnhem-Apeldoorn	56.1	70.7	73.6	73.7
Amsterdam-Purmerend	67.4	77.6	81.2	78.3

[a]Time, distance, signals, and hierarchy labels make up this set.
[b]Initial, not optimal, label coefficient values were applied for Arnhem-Apeldoorn and Amsterdam-Purmerend corridors.
[c]N/A = not applicable.

TABLE 4 ESTIMATION RESULTS FOR REFERENCE SIX-LABEL MODEL

Variable	Coef. Estimate	Standard Error	T-Ratio
Estimated with separate label-specific dummies by area			
Total travel time (minutes)	-0.0979	0.053	-1.8
Total distance (kilometers)	-0.895	0.099	-9.1
Scenic time (minutes)	0.0767	0.025	3.1
Number of traffic signals	-0.138	0.041	-3.4
Expressway distance (km)	0.108	0.022	4.8
High road quality distance (km)	0.365	0.075	4.9
Low road hierarchy time (min)	-0.0877	0.017	-5.3
Overall label-specific dummy variables			
Minimum time route	0.119	0.064	1.9
Minimum distance route	0.881	0.083	10.7
Maximum scenic route	-0.0527	0.088	-0.6
Minimum signals route	0.722	0.081	8.9
Maximum capacity route	0.571	0.077	7.4
Hierarchical travel route	0.888	0.076	11.7

	First Run	Second Run
Total number of observations:	3052	3052
Likelihood with zero coeffs.:	-3459.6	-3459.6
Final Likelihood:	-1038.7	-1363.4
$p^2(0)$:	0.700	0.606

label models resulting from this analysis. These models include seven generic level-of-service variables, and Tables 4 and 5 show the coefficient estimates for all variables included in the alternative utility functions. The label-specific dummy variables take a value of 1 if the indicated label is matched by the route alternative in question and 0 otherwise.

Initial models with generic level-of-service variables and label-specific dummies applicable across study areas failed to yield significant, correctly signed level-of-service coefficients (e.g., total travel time). Thus a two-step estimation process was used to develop the models presented in Tables 4 and 5. First, a model specification with separate sets of label-specific dummies for each area was estimated, producing significant level-of-service coefficient values of correct sign and relative magnitude. Because this model cannot be applied generally with respect to geographic area, a second estimation is required, yielding values for a single set of label-specific dummies while constraining the level-of-service coefficients to the values obtained in the previous estimation.

Evaluation of Choice Models

The choice models of Tables 4 and 5 can be compared in several terms, including data requirements, chosen route coverage, goodness of fit, and values of model coefficients.

TABLE 5 ESTIMATION RESULTS FOR REFERENCE FOUR-LABEL MODEL

Variable	Coef. Estimate	Standard Error	T-Ratio
Estimated with separate label-specific dummies by area			
Total travel time (minutes)	-0.198	0.061	-3.3
Total distance (kilometers)	-0.577	0.101	-5.7
Scenic time (minutes)	0.145	0.031	4.7
Number of traffic signals	-0.0849	0.052	-1.6
Expressway distance (km)	0.0936	0.031	3.0
High road quality distance (km)	0.206	0.072	2.8
Low road hierarchy time (min)	-0.0824	0.018	-4.4
Overall label-specific dummy variables			
Minimum time route	0.448	0.060	7.4
Minimum distance route	1.64	0.115	14.2
Minimum signals route	1.10	0.093	11.8
Hierarchical travel route	1.47	0.098	15.0

	First Run	Second Run
Total number of observations:	2635	2635
Likelihood with zero coeffs.:	-2434.5	-2434.5
Final Likelihood:	-726.5	-1087.8
$p^2(0)$:	0.702	0.553

In terms of data requirements, both models require information for the seven level-of-service attributes for all alternative routes. The only additional difference between the six-label and the four-label models is that the former requires the road capacity data necessary to generate the capacity label. In this analysis, capacity was calculated on the basis of numbers of lanes and road width. The four-label model requires somewhat less computation to run because only four labels must be generated for all relevant origin-destination pairs as opposed to six for the other model.

Considering chosen route coverage, the six-label model was based on approximately 3 percent more (3,667 versus 3,563) chosen route observations than the four-label model. This is because the inclusion of two extra labels in the model specification allowed the analysis of the behavior of an additional sample of drivers to take place—namely, those 104 drivers who were observed to choose a "maximum scenic" or "maximum capacity" route that did not overlap the other four labeled routes.

The likelihood and ρ^2-statistics of each model indicate how well the implied predictions for models about route choice fit the observations for the available sample of drivers. Strictly speaking, the ρ^2-values for these two models are not comparable because they were not estimated on the same set of observations. Nevertheless, keeping these reservations in mind, the ρ^2-statistic of the six-label model apparently indicates somewhat better fit to the data for the two study areas—Arnhem-Apeldoorn and Amsterdam-Purmerend.

The coefficient estimates of the level-of-service variables for both models all have the intuitively correct sign. For example, one would expect increasing travel time to lead to decreasing attractiveness of the alternative, and indeed the travel time coefficient has a negative sign. Similarly, road quality, scenic time, and distance on expressways are all hypothesized as positive qualities of a route, and these variables have positive signs. Another measure that may be used to appraise the reasonableness of a model is its implied "value of time," which is calculated here by determining the ratio between the time and distance coefficients and factoring in an assumed operating cost per unit distance.

Assuming a marginal cost of driving of $0.15 per mile (gasoline costs about $2.90 per U.S. gallon in the Netherlands at current exchange rates), the implied values of time for a major nonexpress road that is neither scenic nor of high quality are $0.61 per hour and $1.92 per hour for the six- and four-label models, respectively. For a minor road that is neither scenic nor of high quality, the respective values are $1.16 per hour and $2.72 per hour. Although all these estimates appear to be on the low side, the values of the four-label models agree more closely with other sources of value-of-time estimates.

In conclusion, the differences between the two final models are not very great. The six-label model (Table 4) is based on more observations and shows better fit to the observed data, whereas the four-label model requires fewer data to operate and has a more reasonable implied value of time. If a choice were to be made between application of one model or the other, the six-label model would be recommended unless capacity data were difficult to come by or the value of time were perceived as too low based on other studies.

In practice, several of the variables used in these models are not likely to be available for the networks to which the models are to be applied. For these circumstances, reduced models were developed in which the requirements for data were substantially reduced or omitted, for example, scenery, road quality, and traffic signals. These models are based on four or even three labels. The loss of explanatory power of these reduced models compared with the models of Tables 4 and 5 is the inevitable consequence of the omission of the relevant variables. Fortunately, some variables other than time and distance, such as hierarchical level, speed limit, and capacity, are generally available in the Netherlands.

Other Results

Apart from the variables incorporated in the models presented in Tables 4 and 5, several other variables were considered for inclusion in the models. Some of these could be eliminated because of their excessively high correlation with variables already included in the models, others because they were not found to significantly influence route choice. In particular, income-dependent effects were carefully tested, but no significant influences could be found.

Further tests were made of differences in behavior among drivers traveling for various purposes, making trips of varying lengths, or traveling with various frequencies. Although some differences of these types were found, they were much smaller than the differences with respect to geographical area.

In general, despite the differences between areas just mentioned, a substantial degree of transferability was found among the three areas for which data were available. As noted above, the labeling procedure was transferable without problems; the choice models lost explanatory power in the transfer but still gave useful and reliable results.

Structural tests were also made on the models estimated. Again some evidence was obtained of failure of the independence assumption on which the logit model is based, but this was not sufficiently serious to cause the structure to be abandoned. Moreover, there was no simple way in which the structural divergence could be approximated.

CONCLUSIONS AND RECOMMENDATIONS

A method has been developed for describing the general route choice behaviors of car drivers. The method is based on the "labeling" of alternative routes that provide realistic possibilities for each driver's journey. A probability model then represents the choice among these route alternatives.

The method is based on a fundamental reassessment of the choice processes that lead to the selection of routes and the analysis of the choices actually made by nearly 7,000 drivers observed in three corridors in the Netherlands.

Several road characteristics other than time and distance are found to be important in influencing route choice. Of particular relevance to policy is the finding that characteristics associated with major roads (restricted access, high speed limit, high capacity, hierarchical status) are strongly positive in influencing route choice. Scenery (positive) and traffic lights (negative) are also found to be relevant.

Even under uncongested circumstances, several routes are used for a given journey. The models estimated identify these routes and predict the proportion of vehicles that will use them. The fact that these predictions are based on models formulated by observing behavior rather than on an arbitrary basis as in some algorithms in current use gives much more confidence in their use.

Application procedures have been developed for the models. These procedures take into account the existing sophisticated methods for the treatment of capacity constraint. The application of the route choice models would add little to the computer time needed to make an assignment and would require little additional software.

Reduced models have been developed to be applied in circumstances of reduced data availability.

Further development of route choice analysis is required to account for two important aspects:

1. The information available to the driver is not currently modeled. Apart from fixed sign posting, interest in formulating policy on the dynamic provision of information is growing, and it is important to know the extent to which drivers might be influenced by methods of providing it.
2. Cost is incorporated into the models only weakly, through the distance variables. The policy under consideration includes "road pricing," whereby drivers would pay much more directly for the use of roads; the influence of such measures on route choice, however, needs to be investigated.

A third aspect that might be considered is the apparent safety of one route compared with another and how that affects route choice.

ACKNOWLEDGMENTS

This project was conducted jointly by Hague Consulting Group and the Transportation Systems Division of the Netherlands

Ministry of Transport (the project's sponsor). The authors wish to acknowledge the staff of the Ministry who contributed to the study. They also wish to thank Theo Bergman, Lionel Silman, Steve Pitschke, and Rohit Ramaswamy, who contributed greatly to the earlier phases of this study.

REFERENCES

1. M. Ben-Akiva, M. J. Bergman, A. J. Daly, and R. Ramaswamy. Modelling Inter-Urban Route Choice Behavior. In *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, Kijkduin, The Netherlands, VNU Science Press, Utrecht, 1984, pp. 299–330.
2. H. Aashtiani and W. Powell. Route Choice Models in the Traffic Assignment Process. Massachusetts Institute of Technology, Cambridge, Mass., 1978.
3. N. Hidano. Driver's Route Choice Model: An Assessment of Residential Traffic Management. Presented at the WCTR, Hamburg, West Germany, 1983.
4. P. H. L. Bovy. Het Kortste-Tijd Routekeuzecriterium: Een Empirische Toetsing. Presented at the Colloquium Vervoerplanologisch Speurwerk, The Hague, The Netherlands, 1979.
5. M. H. Ueberschaer. Choice of Routes on Urban Networks for the Journey to Work. In *Highway Research Record 369*, HRB, National Research Council, Washington, D.C., 1971, pp. 228–238.
6. F. Tagliacozzo and F. Pirzio. Assignment Models and Urban Path Selection Criteria: Results of a Study of the Behavior of Road Users. *Transportation Research*, Vol. 7, 1973, pp. 313–329.
7. C. C. Wright. Some Characteristics of Driver's Route Choice in Westminster. *Proc., PTRC Summer Annual Meeting*, PTRC Education and Research Services, Ltd., London, 1976.
8. J. E. Burrell. Multiple Route Assignment and Its Application to Capacity Restraint. In *Fourth International Symposium on the Theory of Traffic Flow*, Karlsruhe, West Germany (W. Leutzbach and P. Baron, eds.), 1968.
9. R. B. Dial. Probabilistic Assignment: A Multipath Traffic Assignment Model Which Obviates Path Enumeration. Ph.D. dissertation. University of Washington, Seattle, 1970.
10. R. Hamerslag. Onderzoek naar Routekeuze met Behulp van een Gedisaggregeerd Logitmodel. *Verkeerskunde*, No. 8, 1979, pp. 377–382.
11. H. Morisugi, N. Miyatake, and A. Katoh. Measurement of Road User Benefits by Means of a Multi-Attribute Utility Function. *Papers of the Regional Science Association*, Vol. 46, 1981, pp. 31–43.
12. V. E. Outram. Route Choice. *Proc., PTRC Summer Annual Meeting*, PTRC Education and Research Services, Ltd., London, 1976.
13. E. J. Lessieu and J. M. Zupan. River Crossing Travel Choice: The Hudson River Experience. In *Highway Research Record 322*, HRB, National Research Council, Washington, D.C., 1970, pp. 54–67.
14. D. L. Trueblood. Effect of Travel Time and Distance on Freeway Usage. *Bulletin 61*, HRB, National Research Council, Washington, D.C., 1952, pp. 18–35.
15. M. Wachs. Relationship Between Driver's Attitudes Toward Alternative Routes and Driver and Route Characteristics. In *Highway Research Record 197*, HRB, National Research Council, Washington, D.C., 1967, pp. 70–87.
16. R. M. Michaels. The Effect of Expressway Design on Driver Tension Responses. *Public Roads*, Vol. 32, No. 5, 1962.
17. H. J. Wootton, M. P. Ness, and R. S. Burton. Improved Direction Signs and the Benefits for Road Users. *Traffic Engineering and Control*, 1981.
18. M. Ben-Akiva and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, Mass., 1985.
19. Hague Consulting Group. *ALOGIT User Documentation*, The Hague, The Netherlands, 1988.
20. J. D. Bolland, M. D. Hall, and D. van Vliet. SATURN: A Model for the Evaluation of Traffic Management Schemes. *Working Paper 106*. Institute for Transportation Studies, University of Leeds, Leeds, England, 1979.