TRANSPORTATION RESEARCH RECORD
569

New Approaches to Travel Forecasting

10 reports prepared for the 54th Annual Meeting of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
Washington, D. C., 1976

Transportation Research Record 569
Price $6.00
Edited for TRB by Marianne Cox Wilburn

Subject areas
55 traffic measurements
84 urban transportation systems

Transportation Research Board publications are available by ordering directly from the board. They may also be obtained on a regular basis through organizational or individual supporting membership in the board; members or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competence and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in individual papers and attributed to the authors of those papers are those of the authors and do not necessarily reflect the view of the committee, the Transportation Research Board, the National Academy of Sciences, or the sponsors of the project.

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA

National Research Council. Transportation Research Board.
New approaches to travel forecasting.

(Transportation research record; 569)

1. Choice of transportation–Congresses. 2. Traffic estimation–Congresses. 3. Travel–Congresses. I. Title. II. Series.
TE7.H5 no. 569 [HE336.C5] 380.5'08s [380.5'1]
ISBN 0-309-02482-X 76-40324

CONTENTS

FOREWORD

The papers contained in this RECORD focus on new and emerging techniques in travel forecasting that supplement traditional techniques and extend the potentials of forecasting technology.

Difiglio and Reed describe a transit sketch planning process that uses an aggregated data base. A logit modal-choice model is adapted to operate at a level of aggregation that produces transit demand estimates consistent with detailed estimates. Difiglio and Reed state that policy alternatives can be evaluated with respect to various levels of impact, and that, with cost analysis, the degree of subsidy for each policy alternative can be estimated. Policies that appear most promising can be further tested in the traditional urban transportation planning system to evaluate the detailed impacts at zone-to-zone and facility levels.

Brand focuses on research related to improving the understanding of travel behavior. The paper discusses 7 major research issues around which research on travel behavior may be structured. Recent findings related to each question are presented, and current uncertainties and untested hypotheses are exposed for discussion.

Lerman and Ben-Akiva describe a series of disaggregate behavioral models that forecast the probability of a household's selecting various automobile-ownership and mode-to-work combinations. The model assumes that workplace and residential location are predetermined. The paper describes the considerations in choosing independent variables and specifying utility functions. The estimation results for each of 7 distinct socioeconomic groups or market segments with different behavioral distributions are presented and analyzed.

Burns, Golob, and Nicolaidis focus on automobile-ownership behavior that is modeled as a function of socioeconomic factors and the availability and levels of service of public transportation systems. According to the authors, the results from the initial tests of some hypotheses of the automobile-ownership theory are encouraging. Estimated coefficient values of the variables of the models derived for the theory are correctly signed in all cases, and the traditional goodness-of-fit measures are at values that are acceptable for nonlinear estimation equations of the multinominal logit type. In conclusion, Burns, Golob, and Nicolaidis state that the models help to begin to identify causal mechanisms in urban-household travel behavior but that much remains to be done before the models can be effectively applied in predicting automobile-ownership changes that result from transportation system changes.

Smith and Cleveland use data from 1953 and 1965 home-interview surveys in Detroit to test the time stability of disaggregate trip-generation and predistribution modal-choice models. According to Smith and Cleveland, initial cross classification analysis showed 4 to 18 percent increases in household trip generation for households with cars available. A statistical test of the overall time stability of multiple, linear-regression trip-generation equations indicated that the equations were not stable unless non-trip-making households were removed. The individual regression coefficients also were tested for time stability, and, despite the lack of overall statistical time stability, disaggregate work and home-based trip-generation equations for 1953 produced reasonable estimates of 1965 zone-level trips.

Zaryouni and Kannel discuss research results related to trip-distribution functions that may be appropriate for estimating zone-trip interchange in small- to medium-sized urban areas. The proposed model is one in which trip-generation and trip-attraction estimates can be obtained primarily from census data. Although the validity of the distribution model has been tested, the total synthetic modeling approach must still be examined.

Berg, Koushki, Krueger, and Bittner discuss a recreation-travel simulation model developed for use in analyzing the impact of outdoor recreation travel by residents of a 9-state Upper Midwest region to Michigan, Minnesota, and Wisconsin. Travel data were collected for 6,441 randomly selected households by using a telephone home-interview survey procedure.

Ben-Akiva and Richards describe a disaggregate modal-choice model with 6 travel modes. A number of alternative model specifications were tested, and the results of these tests were analyzed. The model specification that was considered to be most satisfactory overall is based on treating in-vehicle travel time as a generic variable and out-of-vehicle travel time as a series of modal-specific variables.

Peers and Bevilacqua discuss a set of direct-demand models for estimating intercity transit travel for a Sacramento-Stockton-San Francisco Bay Area corridor study. A series of judgments are described that identify why structural models were chosen instead of operational models and why direct-demand models were used rather than probabilistic-choice models. The methodology of calibration, including various selection and equation development, validation, and forecasting, is outlined.

Adler and Ben-Akiva discuss a joint frequency-, destination-, and modal-choice model for shopping trips that is an extension of models developed earlier. Estimation of the expanded joint-choice model proved to be feasible and resulted in acceptable parameter values. Adler and Ben-Akiva also give an example of an application of the shopping model.

TRANSIT SKETCH PLANNING PROCEDURES

Carmen Difiglio and Marshall F. Reed, Jr., Highway Users Federation

The current urban transportation planning process is highly dependent on a complex set of travel demand models that operate at a relatively fine level of detail. These models ultimately produce travel assignments on alternative modes of travel. Unfortunately, these travel demand estimates are often insensitive to policy variables, and the process is often too cumbersome and time consuming to be used to test a wide range of transportation policy alternatives. This paper describes a complementary analytical method called the transit sketch planning process. Its purpose is not to supplant traditional, more detailed urban transportation planning but to extend the range of existing procedures. The transit sketch planning process uses a much more aggregated data base. A logit modal-choice model is adapted to operate at such a level of aggregation while producing transit-demand estimates that are consistent with detailed estimates. Policy alternatives can be evaluated with respect to their overall impact on central business district, city center, and suburban transit demand. With a complementary cost analysis, the degree of subsidy required for each policy alternative is estimated. Those policies that appear to be most promising can be further tested in the traditional urban transportation planning system to evaluate the detailed impacts at a zone-to-zone and facility level. The procedures do not attempt to advance the state of the art of transportation demand forecasting. Rather they attempt to use the best of existing procedures in a framework that can provide quick responses to transit policy questions that local planners must answer.

•SEVERAL urban areas, especially those of moderate size, are confronted with difficult planning problems precipitated by the need to plan for public transit. Many urban area planners, in the past, have concentrated their efforts entirely on highway planning and have left public transit issues to the private sector or the public operating agency. Even the largest regional planning agencies that have been involved in planning public transit typically have not developed policy-sensitive procedures for evaluating public transit.

The current urban transportation planning process is highly dependent on a complex set of travel demand models that operate at a relatively fine level of detail. These models ultimately produce travel assignments on alternative modes of travel. Unfortunately, these travel demand estimates are often insensitive to policy variables and the process is often too cumbersome and time consuming to be used to test a wide range of transportation policy alternatives. A complementary analytical method, called the transit sketch planning process, is suggested to fulfill the need for a streamlined, policy-sensitive planning procedure. Its purpose is not to supplant traditional, more detailed urban transportation planning but to extend the range of existing procedures.

At the center of this process (Figure 1) is a model or formula that, when calibrated, synthesizes current ridership from base-year trip information and then is used to forecast future transit ridership. The primary objective of the model is to simulate the potential changes in transit ridership resulting from alternative transit capital and operating policies. Because of this, the model is designed to be sensitive to variables that reflect transportation planning policy such as transit accessibility, transit speed and fares, automobile speed, and parking charges. The other nonpolicy variables used are population, work force, and automobile availability, but these are used only to determine captive transit ridership.

MODAL-SPLIT MODEL

Although several transportation supply variables theoretically could be identified as having an impact on the choice of travel mode, only time and cost of travel have been statistically tested. It certainly would be desirable to evaluate other characteristics of transit, such as travel comfort, but without new data sources these variables cannot be handled quantitatively. Their consideration must be addressed in a more qualitative marketing analysis performed at a more detailed planning level.

All modal-split-model variables used are either travel cost or travel time. Travel cost includes the out-of-pocket automobile costs for gas, oil, and parking charges and transit fares for transit travel.

Travel time for the transit mode is divided into 2 types: (a) access and wait time and (b) line-haul (on-board) time. These are converted into monetary values by using value-of-time estimates that are substantially different for each type of time spent in traveling. Automobile travel time also converts to a monetary value.

The difference between the total cost of transit travel and automobile travel is used to determine the percentage of noncaptive trip makers who select the transit mode. A nonlinear modal diversion curve is used for this purpose:

$$P_i = \frac{1}{1 + e^{-aX_i}} \qquad (1)$$

where

P_i = probability of transit mode choice for trip purpose i,
a = calibration factor, and
X_i = difference between automobile and transit cost for purpose i (equation 2).

$$X_i = C_{hwy} - C_{tr} - W_{Ai} A_{tr} + W_{Li} (L_{hwy} - L_{tr}) \qquad (2)$$

where

C_{hwy} = out-of-pocket automobile costs = D at \$0.05/mile (\$0.03/km) + parking + tolls where D = typical trip distance,
C_{tr} = transit fare,
W_{Ai} = value of wait or walk time for each trip purpose i,
A_{tr} = accessibility time to and from transit service (includes half headway time),
W_{Li} = value of vehicle trip time for each purpose i,
L_{hwy} = automobile trip time, and
L_{tr} = line-haul trip time on transit.

The modal-split formula is a logit probability model. When graphed as a function of its independent variables, X_i, the resulting diversion curve appears as shown in Figure 2. It should be noted that the possible values of the dependent variable (probability of modal choice) range between 0 and 1. This formula has certain properties that are consistent with common sense or intuitive notions regarding modal choice. For example, a cost differential of 0 between automobile and transit ($X_i = 0$) implies an equal split between automobile and transit use ($P_i = 0.5$). Also, at this point where transit and automobile are equally attractive, changes in either automobile or transit costs will have the largest impact on modal use. At the point on the diversion curve where $P_i = 0.5$ and $X_i = 0$, the curve has the steepest slope.

The slope of the diversion curve is quite flat at the extreme values of X_t. When transit service is poor relative to automobile travel, the variable takes on a large negative value. In that range, improvements in transit service bring about a very small increase in transit diversion. Similarly, if transit service is far better than automobile service (X_t has a large positive value), improvements in transit service also bring about a small increase in transit diversion.

Because the logit function is asymptotic to the extreme probabilities 0 and 1, the estimated probability of transit usage is never 0 and 1. Even under the most extreme cost differentials, some individuals may choose to take the most costly mode even though the probability of such a choice is low.

MODEL CALIBRATION

The modal-split model can be calibrated by using a trial-and-error process instead of the usual approach of using regression analysis when thousands of travel interchanges are involved. Only 3 parameters need to be determined:

1. W_A, the value of wait or walk time,
2. W_L, the value of in-vehicle time, and
3. a, the exponential factor in equation 1.

In all of the applications of the model thus far, the exponential factor a has been set equal to 2. This factor can be interpreted as determining the spread of the diversion curve in Figure 2 along the x axis. A high value of a produces a diversion curve that is quite steep, which indicates a great sensitivity of modal choice to the differences in transit and automobile trip costs. A low value of a produces a gradual diversion curve that shows a relative insensitivity of modal choice to travel cost differences. These effects are shown in Figure 3.

The value-of-time parameters W_A and W_L are unknown but are reasonably bounded by the research already available on the value of traveler's time. Often this research determines the value of time as a function of the traveler's income. However, because of the large areal aggregations used in this sketch-planning process and the uncertainty involved in forecasting future income levels, income was not used explicitly to set the value of time.

Initial trials for the value-of-time parameters used \$3.00/h for walk and wait time and \$1.00/h for in-vehicle time. The final value-of-time parameters selected differed only slightly from these levels.

Calibration Process

The overall process of calibrating the model to estimate base-year transit trips is illustrated by using work-trip data for Nashville, Tennessee.

The Nashville urban area was divided into 3 sectors (Figure 4).

1. Sector A is the central business district (CBD).
2. Sector B is the remainder of the Nashville city center.
3. Sector C is the suburbs.

A 9-cell trip table showing the number of base-year trips made between and within all sectors was assembled from home-interview travel survey data for the home-to-work and home-to-nonwork trip purposes. The base-year home-to-work trip tables for total trips and transit trips are given in columns 2 and 3 of Table 1.

The modal-split model estimates the probability of choosing the transit mode over the automobile mode. It presumes that there is a choice of mode, but many individuals do not have access to the automobile mode for many of their trips and are captive transit riders. To properly calibrate the modal-split model, one should subtract captive

Figure 1. Transit sketch planning process.

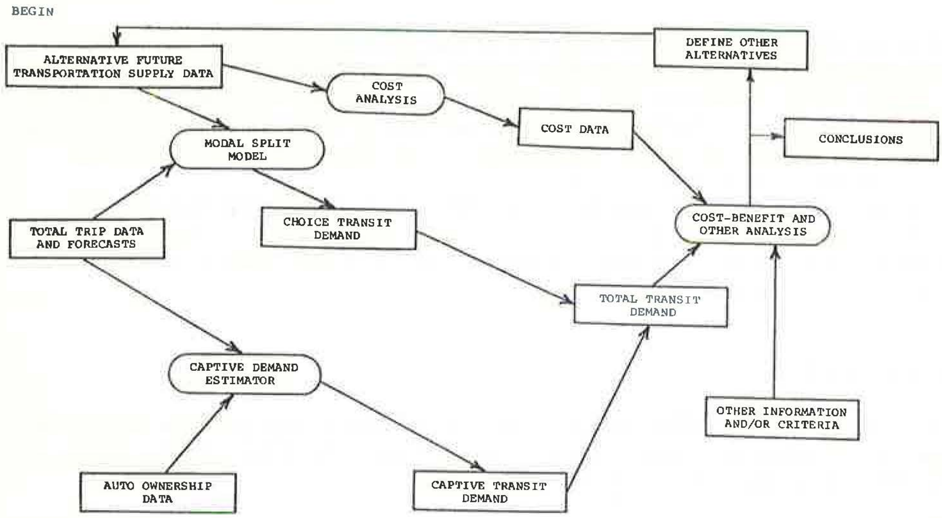


Figure 2. Diversion curve.

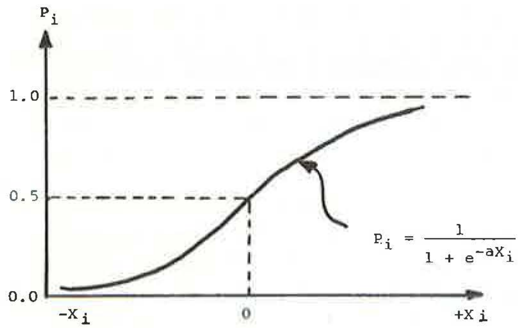


Figure 3. Diversion curve spread.

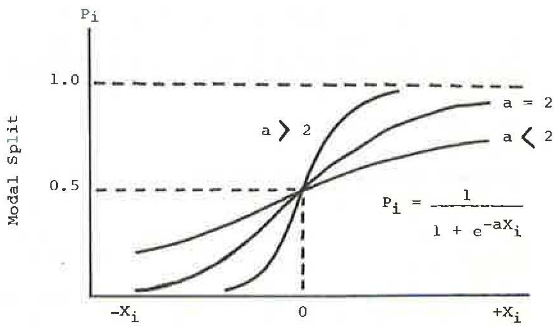
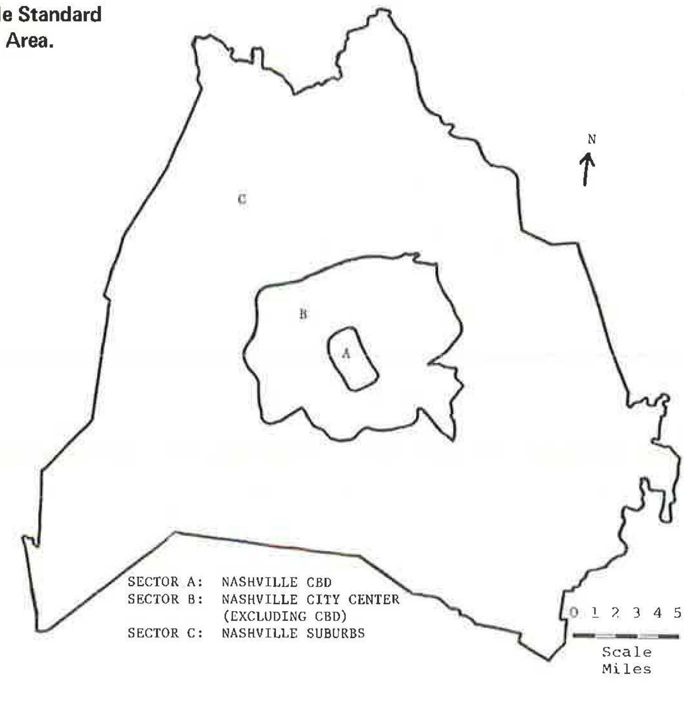


Figure 4. 1960 Nashville Standard Metropolitan Statistical Area.



rider trips from observed total transit trips. Unfortunately, adequate survey data on captive ridership are not generally available. But, it is widely acknowledged that captive ridership is, in many instances, a significant proportion of total transit patronage, especially when transit service is unattractive. An estimate of captive ridership, however compromised by the lack of reliable data, is superior to no estimate at all. No estimate is implicitly an estimate of no captive ridership. A 0 estimate would likely bias the calibration of the modal-split model in low-income trip corridors.

The number of captive transit travelers making a work trip in each sector was estimated by analyzing automobile ownership and work-force data. The excess number of work-trip origins over automobile seats available for the work trip was determined by using estimates of the automobile occupancy rate and the proportion of automobiles used for the work trip.

Having determined the number of captive transit trip origins in each sector, we distributed the trip destinations by using information regarding the commuting patterns in Nashville by socioeconomic class in the base year. This information was not quantitative but rather was based on the predominance of domestic workers traveling between certain sectors. The resulting captive transit demand trip table is given in column 4 of Table 1.

The model variables for each of the 9 trip interchanges were aggregated from base-year highway and transit "networks" for the Nashville area. Each model variable was averaged over all zone pairs in each sector-to-sector interchange. The model inputs for Nashville are given in Table 2.

Aggregation Bias

Typically, modal-split models are used in the urban transportation planning process to determine the percentage of transit use between pairs of zones. The zone is a very small area compared to the sectors used in this sketch-planning process.

It is desirable for the modal-split model to be accurate regardless of the level of areal aggregation. Consequently, the modal-split model was calibrated by using both zone-to-zone travel characteristics and sector-to-sector travel characteristics. However, because of the nonlinear nature of the modal-split model, the forecasted sector-to-sector modal split tends to be biased downward relative to the modal splits calculated at the zone-to-zone level.

This effect is especially large for suburban trips because the degree of variation of suburban transit service is very high, and many areas have no transit service at all.

Fortunately, accurate estimation of the degree of bias from available Nashville transit data was possible. Each input variable in Table 2 was aggregated from zone-to-zone travel characteristics. For example, in sector A to sector C, the average transit line-haul time was 24 min. Of the zone pair interchanges that comprise the sector A to sector C interchange, the transit line-haul times could vary substantially. The degree of variation can be measured by the standard deviation σ. The σ of each input variable therefore was determined.

Calculating the amount of aggregation bias in any particular instance by using modal-split model calibrations at both the zone-to-zone and sector-to-sector level is also possible. The calculated bias (expressed as a ratio) was regressed with the σ of all model variables. The ratio was calculated for individual observations of the regression by performing a number of individual zone-to-zone modal-split estimates and an aggregate estimate based on the average zone-to-zone characteristics in Nashville. Because of the relatively large variance of the variable for transit access (wait and egress time), explaining 96 percent of the variance of the bias ratio by using only the standard deviation of that variable was possible.

$$\frac{\hat{\hat{P}}}{\hat{P}} = 0.41656 + 0.09530(\sigma_{A_{tr}}) \tag{3}$$

for all $\sigma_{A_{tr}}$ such that 6.12214 $\sigma_{A_{tr}} < 30$ and the coefficient of determination $r^2 = 0.962$ where

$\hat{\hat{P}}$ = average of zone-to-zone modal-split estimates selected at random (observation i),

$\hat{P}$ = aggregated zone-group-to-zone-group modal-split estimate—all independent variables used to estimate P_i were averaged to a single set of independent variables, and

$\sigma_{A_{tr}}$ = standard deviation of the independent variable $\sigma_{A_{tr}}$ used in the modal-split model that produced both $\hat{\hat{P}}$ and $\hat{P}$.

An alternate aggregation correction equation, also based on the variance of the logit model variables, is:

$$\hat{\hat{P}}_i = \hat{P}_i + [\mathrm{Var}(X_{iz})]P_i(P_i - 1)(P_i - 0.5) \tag{4}$$

where

$\hat{\hat{P}}_i$ = expected value of P_i aggregated from zone-to-zone modal-choice estimates;

$\hat{P}_i$ = value of P_i evaluated at the average sector-to-sector travel characteristics for automobile and transit—the mean values of all model variables are used to calculate X_i, which is used to determine P_i; and

$\mathrm{Var}(X_{iz})$ = variance of X_i for each zone-to-zone interchange—the function X_{iz} would be used to calculate an individual modal choice for zone pair z.

This formulation is an approximation based on a Taylor expansion of the logit estimator of P_i about the mean of X_{iz} provided by Talvitie (18). It requires estimating a more complex variance than does the regression approach, but the estimates can be used over any range of P_i, and the regression relationship can be applied only over a range of P_i within that used to estimate parameters. The variance of X_{iz} can be more easily calculated by assuming that the model variables are uncorrelated (18) so that

$$\mathrm{Var}(X_{iz}) = \mathrm{Var}(C_{hwy}) - \mathrm{Var}(C_{tr}) - \{(W_{Ai})^2[\mathrm{Var}(A_{tr})]\} + \{(W_{Li})^2[\mathrm{Var}(L_{tr})]\} \tag{5}$$

Nashville Modeling Results

After several trial-and-error iterations to account for the aggregation bias and several estimates for time values, the modal-split model was calibrated to produce base-year choice transit ridership within acceptable limits. The estimates of modal split and choice transit demand are given in columns 5 and 6 of Table 1. The estimated choice transit demand plus the estimated captive transit riders (column 4 in Table 1) produces a total transit demand estimate that can be compared with observed transit demand in the base year and in column 7 in Table 1. The error of the estimated versus the observed transit demand is best expressed as a percentage and an absolute number because either may be misleading unless total trips are taken into account. These errors are given in column 8 and column 9 in Table 1.

SELECTING TEST ALTERNATIVES

Ridership tests of alternative systems are a relatively simple and quick procedure.

Therefore, depending on the scope and the time limitations of the transit planning project, a great many alternatives can be analyzed at least for their impact on ridership. However, because of constraints on time and personnel to develop cost estimates, the number of test plans subjected to the complete analysis should be limited to those that are likely to show discernible differences in either ridership or cost effectiveness.

In designing alternative test systems, the number of changes to transit service variables should be limited. Ideally, only 1 variable at a time should be changed in each test to fully assess the effects of the change. Both significant reductions in fares and increases in line-haul speed will enhance ridership, but taking both into account in 1 test will make it difficult to assess the cost effectiveness of either service change.

A likely candidate for the initial test of the future of transit is the extension of current or base-year service levels (accessibility, speed, and fares) to target-year populations and land areas. This initial test serves as a base for comparing ridership, costs, and benefits of service improvements in other alternative plans. This was identified as alternative A for forecasts that were based on the Nashville application of this sketch-planning method.

Because improved accessibility (the average time to get to the rail station or bus stop plus the average wait time for a transit vehicle) has been found by transit planners and researchers to be the most important factor in enhancing ridership, a basic test alternative is one that improves accessibility. In physical terms, improved accessibility means either increased route miles (kilometers) or decreased headways or both.

Cutting transit-vehicle headways in half will cut the average wait time for patrons. Wait time, because of its problem in inclement weather and its general uncertainty, has a significant influence on transit ridership. This improvement means a doubling of transit vehicles in service. This was identified as alternative D; a more modest reduction of headways by a third was identified as alternative B.

New freeway express-service corridors were selected in appropriate Nashville corridors from sector C to sector A. On the basis of the available freeway corridors, 30 percent of the suburban market was assumed to be served by the express bus service. Alternatives C and E were then defined to be alternatives B and D respectively with express bus service added.

In general, the current bus or rail transit systems in the large urban areas provide regularly scheduled service to all parts of the city center but relatively little service in suburbs. The density of city center development makes scheduled service efficient in the city center but inefficient in the lower density suburbs. To narrow the suburb and city center difference in transit service levels, demand-responsive, small-bus service was added in the suburbs of each test plan. The dial-a-ride service included in the test plans closely approximates existing systems in the United States and Canada. Some of the dial-a-ride trips generated are merely a short link in a longer trip in which line-haul bus or rail service is the principal service mode. These are accounted for in the ridership estimates for the principal mode. Other dial-a-ride trips were principally nonwork, suburb-to-suburb trips and were estimated to attract 2 percent of the total suburb-to-suburb trips. The modal-choice model was not used to estimate these trips because the wait time characteristics of this type of service cannot be compared with scheduled, fixed-route service.

Fare policy also could be tested by using additional alternatives. However, for the purposes of this exercise in applying the sketch planning procedure, fares were not changed. Many studies (2, 5, 10, 11, 12, 13, 14, 17, 19, 20) have shown that reducing fares has less significant influence on transit ridership than other service improvements have. The resulting fiscal deficit from fare reduction (caused by a relatively low demand elasticity of transit fare) requires additional funding that might better be spent on further service improvements to gain ridership. Consequently, fares were held at the calibration-year level, but we emphasize that the sketch planning method we have described can quite easily account for different future-year fare policies.

The 5 test alternatives can be summarized as follows:

1. Service for alternative A is the same as 1959 service but with routes extended

to serve the expanded urban area and population;

2. Service for alternative B is the same as for alternative A but with a 33⅓ percent decrease in bus headways, a 50 percent increase in buses, and dial-a-ride service added throughout most of the suburbs;

3. Service for alternative C is the same as for alternative B but with freeway express bus service;

4. Service for alternative D is the same as for alternative A but with a 50 percent decrease in bus headways, 100 percent increase in buses, and dial-a-ride service added throughout most of the suburbs; and

5. Service for alternative E is the same as for alternative D but with freeway express bus service.

Transit fares for all 1990 alternatives equal 1959 levels adjusted to 1970 dollars.

NASHVILLE STUDY RESULTS

The results of the travel demand analysis for the 5 alternative plans are given in Table 3. The largest increases of ridership result from wait time reductions provided by alternatives B and D. The express bus service provided in alternatives C and E does not change the overall ridership estimates as significantly, but the proportion of suburb-to-center-city riders does increase quite significantly (Table 4).

In Table 4, ridership by each sector-to-sector interchange by available mode of travel is given for each forecast alternative. Because work trips and nonwork trips were estimated separately (peak and off-peak travel characteristics were different in the base year), the impact of each alternative on peak-hour CBD-oriented travel can be identified as follows:

	Average Daily Work Trips to and From CBD		
Alternative	Transit	Automobile	Percent Transit
A	21,713	98,287	19
B	25,073	94,927	21
C	28,877	91,123	24
D	27,599	92,401	23
E	30,977	69,023	26

No attempt is made here to analyze the competing alternatives and make any recommendations. The numbers produced show the kind of planning data yielded by the sketch planning process. Other considerations, including cost and income and, ultimately, detailed network-level testing, should be taken into account before specific public transit programs are implemented.

The sketch-planning process does not account for various qualitative changes in transit service that should be evaluated (air conditioning, carpeting, and reduced noise). Also any directly related financial analysis would not necessarily account for possible improvements in operating efficiency of transit operators. Similarly, proposals to deregulate the urban transit industry to allow paratransit competition (jitneys) with current scheduled bus and regulated taxicab service must be evaluated outside the context of the described procedures.

Table 1. Actual and estimated trip demand.

Sector-to-Sector Trip Interchange (1)	Total Demand[a] (2)	Observed Transit Demand[a] (3)	Captive Transit Demand Estimate (4)	Modal-Split Estimate (5)	Choice Transit Demand Estimate (6)	Choice + Captive Transit Demand Estimate (7)	Percentage Error of Estimated Versus Observed Transit Demand (8)	Absolute Overestimate or Underestimate of Observed Transit Demand (9)
A-A	1,205	232	0	0.325	392	392	+69.0	+160
A-B	1,337	232	34	0.150	201	235	+1.3	+3
A-C	473	162	137	0.054	26	163	+0.6	+1
B-A	25,898	6,095	313	0.225	5,827	6,140	+0.7	+45
B-B	29,953	3,707	1,251	0.083	2,486	3,737	+0.8	+30
B-C	10,547	1,861	1,564	0.016	167	1,731	-8.5	-166
C-A	13,256	1,144	0	0.085	1,127	1,127	-1.5	-17
C-B	12,621	266	0	0.016	202	202	-10.6	-24
C-C	13,677	80	0	0.005	68	68	-15.0	-12
Total	108,967	13,769	3,299		10,496	13,795	0	+26

[a]From 1959 Nashville home-interview survey, internal trip file.

Table 2. 1959 work-trip characteristics.

Sector-to-Sector Trip Interchange	Transit Fare (cents)	Transit Access[a] (min)	Transit Line-Haul Time (min)	Highway Time (min)	Parking Cost (cents)
A-A	15	8	8	4.4	24
A-B	15	15	15	9.3	0
A-C	20	32	24	17.0	0
B-A	15	15	15	9.3	24
B-B	15	20	24	11.6	0
B-C	20	41	34	19.0	0
C-A	20	32	24	17.0	24
C-B	20	41	34	19.0	0
C-C	15	45	25	14.0	0

[a]Wait and egress time.

Table 3. 1990 average daily work trips on transit.

Alternative	Freeway Bus	Bus	Demand Responsive	Total	Modal Split	Increase Over Plan A (percent)
A	–	57,587	–	57,587	5.2	–
B	–	75,627	6,920	82,547	7.5	43
C	5,352	74,079	6,920	86,351	7.9	51
D	–	92,038	6,920	98,958	9.0	72
E	5,352	90,070	6,920	102,342	9.7	77

Table 4. 1990 average daily passenger trips.

Sector-to-Sector Trip Interchange	Service	Alternative A	Alternative B	Alternative C	Alternative D	Alternative E
A-A	Bus	2,416	2,668	2,668	2,796	2,796
A-B	Bus	1,254	1,433	1,433	1,547	1,547
A-C	Bus	291	391	391	485	485
B-A	Bus	27,337	30,425	30,425	32,457	32,457
B-B	Bus	16,321	25,001	25,001	31,361	31,361
B-C	Bus	2,894	4,544	4,544	6,009	6,009
C-A	Bus	4,372	7,131	5,583	9,503	7,535
C-A	Express bus	0	0	5,352	0	5,352
C-B	Bus	1,248	2,848	2,848	4,272	4,272
C-C	Bus	1,454	1,186	1,186	3,608	3,608
C-C	Dial-a-ride	0	6,920	6,920	6,920	2,676
Total		57,587	82,547	86,351	98,958	102,342

SKETCH PLANNING PERSPECTIVE

The principal purpose of this transit sketch planning process is not to determine the future with precision but to compare the probable impact of alternative transit planning policies. We understand that relationships such as those used in the modal-split model may explain a given set of data for a particular time period but may not always hold true when used to forecast future activity. Even with the most sophisticated procedures, cause-and-effect relationships between current and future behavioral data cannot be determined. The future is always full of imponderables—and indefinitely numerous sets of possible courses of development—that never can be determined with certainty.

More sophisticated modeling procedures are certainly possible, and some are being advanced in many large, urban, regional transportation planning studies. However, the sketch planning process described here serves a different role, one in which alternative transit policies can be quickly evaluated, which is a task that no detailed urban modeling process has yet accomplished.

REFERENCES

1. 1972-1973 Transit Fact Book. American Transit Association, Washington, D.C.
2. M. Beesley. The Value of Time Spent in Travelling: Some New Evidence. Economica, Vol. 32, 1965.
3. D. Brand. Travel Demand Forecasting: Some Foundations and a Review. HRB Special Rept. 143, 1973, pp. 239-282.
4. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, 1972; NTIS, PB 210 515.
5. Charles River Associates. An Evaluation of Free Transit Service. U.S. Department of Transportation, 1968; NTIS, PB 179 845.
6. T. A. Domencich, G. Kraft, and J.-P. Vallette. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968, pp. 64-78.
7. M. J. Fertal et al. Modal Split. U.S. Bureau of Public Roads, 1966.
8. Economic Characteristic of the Urban Public Transportation Industry. Institute for Defense Analyses, Arlington, Va., Feb. 1972.
9. E. J. Kannel and K. W. Heathington. Temporal Stability of Trip Generation Relations. Highway Research Record 472, 1973, pp. 17-27.
10. M. A. Kemp. Some Evidence of Transit Demand Elasticities. Urban Institute, Washington, D.C., 1973.
11. M. A. Kemp. Reduced Fare and Fare-Free Urban Transit Services—Some Case Studies. Urban Institute, Washington, D.C., 1973.
12. M. A. Kemp. Transit Improvements in Atlanta—The Effects of Fare and Service Changes. Urban Institute, Washington, D.C., 1974.
13. C. A. Lave. A Behavioral Approach to Modal Split Forecasting. Transportation Research, Vol. 3, No. 4, 1969.
14. T. E. Lisco. The Value of Commuters' Travel Time—A Study in Urban Transportation. Department of Economics, Univ. of Chicago, PhD dissertation, 1967.
15. B. V. Martin, F. W. Memmott, and A. Bone. Principles and Techniques of Predicting Future Demand for Urban Area Transportation. M.I.T., Rept. 3, 1964.
16. J. R. Meyer, J. F. Kain, and M. Wohl. Urban Transportation Problem. Harvard Univ. Press, Cambridge, Mass., 1966.
17. P. R. Stopher. A Probability Model of Travel Mode Choice for the Work Journey. Highway Research Record 283, 1969, pp. 57-65.
18. A. Talvitie. Aggregate Travel Demand Analysis With Disaggregate or Aggregate Travel Demand Models. Transportation Research Forum, Vol. 14, No. 1, 1974.
19. T. C. Thomas. Value of Time for Commuting Motorists. Highway Research Record 245, 1968, pp. 17-35.
20. T. C. Thomas and G. I. Thompson. The Value of Time for Commuting Motorists

as a Function of Their Income Level and Amount of Time Saved. Highway Research Record 314, 1970, pp. 1-19.

21. A. R. Tomazinis. Modal Split Model in the Penn Jersey Transportation Study Area. Highway Research Record 165, 1967, pp. 41-75.
22. A. R. Tomazinis. Forecasting Travel Demand for New Urban Transportation Systems. Highway Research Record 392, 1972, pp. 26-35.
23. 1972 National Transportation Report. U.S. Department of Transportation, July 1972.
24. Los Angeles Metropolitan Area Mode Choice Model Development Study. Alan M. Voorhees and Associates, 1972.
25. S. C. Warner. Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice. Northwestern Univ. Press, Evanston, Ill., 1962.
26. M. Wohl and B. V. Martin. Traffic Systems Analysis for Engineers and Planners. McGraw-Hill Book Co., New York, 1967.
27. R. J. Wallin and P. H. Wright. Factors Which Influence Modal Choice. Traffic Quarterly, Vol. 28, No. 2, 1974.

APPROACHES TO TRAVEL BEHAVIOR RESEARCH

Daniel Brand, Department of City and Regional Planning, Harvard University

Research approaches to improving the understanding of travel behavior are outlined in this paper. The objective of the travel behavior research is to make more informed judgments concerning trade-offs between the basis in behavior, and thus the logic and plausibility of travel forecasts, and the time, money, and skills required to carry out the forecasts. The behavioral assumptions to be evaluated relate to the perceptions, valuation, and structure of choice and possible conditioned or learned behavior resulting from the stimuli that give rise to travel decisions for various household and individual travelers. The paper focuses on 7 major issues around which research on travel behavior may be structured. Some recent findings related to each issue are presented, and current uncertainties and untested hypotheses are exposed for discussion. The significance of the research issues also is discussed. Criteria for evaluating alternative behavioral approaches to travel theory are presented.

•THE PURPOSE of this paper is to describe various travel behavior research approaches for evaluating the behavioral assumptions underlying the different models currently in use and models being proposed for passenger travel demand forecasting. The behavioral assumptions to be evaluated relate to the perception, valuation, structure of choice, and possible conditioned or learned behavior resulting from the stimuli that give rise to travel decisions for various households and individual travelers.

In transportation planning, planners have come to take for granted the detailed simulations and forecasts of travel on multimodal transportation networks that current urban travel estimating procedures make possible. Social scientists are awed by the boldness of planners in applying at different times and places relationships derived by using data from choices exercised over limited sets of travel and household-location opportunities. Psychology, for example, is much more tentative and fundamental. Models of perception and preference are derived from precisely defined axioms. Behavior is observed under controlled and fully described conditions. Burnett (27) said, "What knowledge we possess of the relations between human learning, perception, and choice derives from experimental psychology, and the connections between the simple non-spatially-oriented decision-making tasks of the laboratory and real life spatially oriented decision-making are tenuous." Travel modelers working on transportation studies have not enjoyed such luxury.

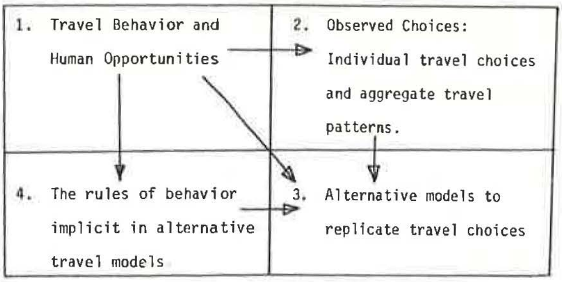
Nevertheless, transportation planning has made sizable strides in recent years in applying a variety of travel models that replicate urban travel patterns. The models address a great variety of transportation policy issues arising at both the aggregate policy level (for example, free transit) and the more fine-grained network levels of detail.

Recently, travel modelers have analytically derived several of the rules of travel behavior implied (or required) by the alternative travel modeling approaches that have been taken (21). These rules allow them to evaluate how well alternative travel models match their understanding of travel behavior. What travel modelers lack, of course, is the basic understanding of travel behavior.

TRAVEL MODELING

Except for the arrows from box 1, the current state of understanding of the interactions

Figure 1. Travel modeling diagram.



of travel behavior, travel models, and observed travel data may be represented as shown in Figure 1. In Figure 1, travel behavior and the set of spatially distributed human opportunities (box 1) drive the travel choices observed in urban areas (box 2). These, together with critical criteria such as time, budget, and available work-force skills, drive the travel models (box 3) that should be used in a particular transportation planning situation. Recently acquired knowledge of the choice behavior implicit in alternative travel models (box 4) allows use of this information in selecting and using alternative travel models. Notice, however, that no arrows go from the remaining boxes back to the travel behavior box. Statistics cannot demonstrate causality. In this case, statistics cannot discern the reasons why particular travel choices have been made. New approaches to understanding travel behavior are needed. These approaches will have to come from outside the traditional data collection and statistical methods in which transportation planners and researchers become quite expert. Because data can be used ultimately only to reject hypotheses, powerful theoretical development accompanied by careful and informed data collection procedures clearly is needed. Skills and techniques will be required from disciplines that have developed experience in understanding human motivation and behavior.

As transportation planners and economists move toward psychology in trying to understand travel-choice behavior, they find psychologists and other behavioral scientists moving to meet them in exploring ways to mesh individual-level data with aggregate-level data in real-choice situations. The main focuses to date have been where aggregate-level data were readily available (primarily in the fields of election forecasting and market research), but the trend is also apparent in other substantive areas such as demographic research that uses family-planning attitudes and behaviors in various cultures. These developments have become possible only recently mainly because of the increasing availability of high-speed, large-capacity computers and problem-oriented software that implements many of the new psychometric techniques.

The ultimate objective of the research approaches outlined in this paper is to allow trade-offs to be made in specific planning situations between the basis in behavior (and thus the logic and plausibility of travel forecasts) and the time, money, and skills required to carry out the forecasts. For certain kinds of situations, some models will dominate. That is, they will be better on all counts and thus should be used. In effect, a certain kind of code or set of recommendations of "what models to use when" must be a pragmatic and hopefully not naive goal of travel behavior research. Such recommendations and modeling choices always are being made, implicitly or explicitly, based on one's knowledge of all factors in a particular planning situation. However, because of the great variety of planning circumstances, a complete code of recommendations is obviously an elusive goal.

TRAVEL BEHAVIOR

Travel behavior may be defined as the observable reactions of people when they are confronted with choices involving travel. Travel choices include staying in one place or acquiring for the long term either the means of transportation (a car or a prepaid transit pass) or the spatial prerequisites for travel (residence and employment locations). Travel observed in the aggregate is the summation of all individual decisions to locate and travel at a particular frequency and time of day by a particular mode and route to a particular destination.

Clearly, a vast number of variables could be used to describe the array of individual travel options and individual and household opportunity and need structures. The circumstances that give rise to travel and changes in travel behavior are as varied as life itself. An improved understanding of these circumstances (trip purpose, the proposed change in the travel-choice environment, physical and nonphysical characteristics of households, individual traveler characteristics, and length of time at place of residence) requires a systematic and well-structured research program. In this paper, such a research program is structured around a list of research problems.

RESEARCH PROBLEMS

There are 7 research problems or issues.

1. The scaling-specification problem asks the question, What attributes of the travel-choice environment give rise to or influence amount of travel?
2. The choice-abstract versus choice-specific problem, which is a subset of the scaling-specification issue, asks the question, Are the attributes influencing travel, such as modes used to produce transportation, perceived together with travel choice or independently of travel choice?
3. The measurement problem asks the questions, Are perceived values of attributes that influence travel related to "objectively measured" values and, if so, how?
4. The separability issue asks, What is the structure, or set, of alternative travel choices from which the traveler actually chooses; are travel decisions decomposed; and are there "natural" partitionings such as by frequency, time of day, destination, mode, and path?
5. The independent-choice issue asks, Do attributes influencing travel choices vary in their relative and absolute effect from one travel choice, or subchoice, to another, such as from modal choice to destination choice in a presumed hierarchy of subchoices to make up a complete travel decision?
6. The stochastic issue asks, Can travel be considered to be the manifestation of a set of conditioned behaviors that involves learning and changes in behavior over time?
7. The household behavioral unit and competing demands issue asks, What are the circumstances of households, or other basic behavioral units to be identified, that affect the findings in items 1 through 6?

These 7 research approaches together constitute an initial attempt to list the issues relevant to the study of travel behavior. The questions go beyond the scaling of the kinds of stimuli that govern travel behavior, their relative and absolute importance, and their interrelationships (problems 1 and 2) and include questions relating to measurement (problem 3), choice structure (problems 4 and 5), learning and conditioning (problem 6), and the influence of the household and other decision units on travel behavior (problem 7).

Each of these 7 research problems is important in its own right. All of the problems, of course, are interrelated. The first 5 are pragmatically stated questions brought about by the requirements of the travel models in use today or proposed for use in travel forecasting. The last 2 describe more fundamental approaches to understanding human behavior. The following sections outline the substance and significance of each of these research problems.

Problem 1: Scaling

A general statement of the scaling problem is that it is the problem of defining the dimensions of the travel-choice environment that influence travel behavior. The significance of the scaling problem is certainly well recognized. The problem of identifying the important stimuli that influence the selection of travel choices is central to a better understanding of travel behavior. The next few paragraphs indicate how the problem of scaling intersects with other research questions and how its "solution" may ultimately be affected by data limitations.

Definition of Behavioral Variables

Long lists of attributes that may influence travel abound in the literature. However, travel times and costs are behavioral variables just as are comfort, convenience, reliability, safety, weather, number of transfers, and journey units. A usable behavioral model is not one that identifies a new behavioral variable but does not characterize it well enough for it to be measured (problem 3) and forecast. In fact, many definitions of comfort and convenience in transportation include time and cost components that are measured relatively easily (162).

Path analysis (151), a technique for rejecting on the basis of observed behavior certain causal orderings in an assumed completely recursive travel model structure, would appear to hold some promise for defining independent behavioral variables. However, it has seen only limited application in travel demand modeling (88).

The last 10 years have seen a sudden proliferation of psychometric analysis techniques for dealing with data on perception and preferences. At first, psychometric analysis was used primarily in academic psychology, but, more recently, it has come into favor in market research. The techniques of psychometric analysis differ in the ways perceived similarities are represented and in the ways they are mapped into behavior in particular choice situations. Of substantial interest in pursuing the scaling problem are techniques that allow the mapping of aggregate or individual "psychological domains" of items in multidimensional spaces from data on perceived similarities, characteristics, preferences for items, and substitutability of items (67, 106, 185). The initial applications of psychological scaling techniques in travel research are very recent (44, 135, 162) and certainly do not exhaust the capabilities of the technique. Multidimensional scaling is an important research technique that will see increased application in travel behavior research in the near future.

Limits on Human Discrimination Among Attributes

The number of variables that should be included in models of travel behavior must be limited to the number among which humans can discriminate. Theories of human discrimination and choice indicate the presence of a minimum variance needed for discrimination and that the number of variables that can be considered in any single decision is finite (160). One of the contributing factors to making destination choice the most difficult of the short-term travel choices to model is the great variety of alternative destinations that are available and the subtleness and temporal nature of the attributes needed to differentiate among possible destinations (70). Also, it is unclear whether the traveler, when confronted with a large number of travel choices, breaks down the choice into smaller subchoices (problem 4) each of which is characterized by a larger number of attributes, or whether all choices are considered simultaneously but in terms of a smaller number of attributes.

Deducing Values From Behavior

The major emphasis to date in travel-choice modeling has been in deducing the relative importance of modal attributes through people's revealed preferences. Revealed preferences are what one may deduce about people's preferences or values from their behavior. The values are deduced by relating people's observed travel, either as individuals or in the aggregate, to measured changes in transportation system characteristics. For example, the well-known value-of-time studies (108, 167) generally relate travel behavior in a simple model to measurement values of travel time and cost by alternate modes and routes.

However, deducing values from travel behavior is not entirely satisfying because travel choices depend not only on people's values but also on the travel and other opportunities open to them. Boulding (23) said that choices are "necessitated when the elements in the set of choices are scarce, in the sense that there is a limitation in the quantities that can be obtained which prevents the chooser [from] reaching the point of satiety" and revealing his or her true values or preferences.

Value-of-time studies deduce activity-specific, time-cost trade-offs for the limited set of choices (often only 2) that have been defined for the particular study. In reality, travel choices, although limited, are not nearly so restricted. In this sense, value of-time models are in danger of being seriously misspecified. Lansing, Mueller, and Barth (102) stated that "the fact [is] that there is a tradeoff and what that tradeoff is, depends on where you are. Thus, what is a good thing at one point becomes a bad thing at another, and vice versa."

Also, in deducing values from behavior, there is the problem of the limited available measurements of transportation system characteristics. A severe limitation on the use of the most widespread available data on observed travel, namely existing home-interview survey data, is the lack of data on, for example, alternative travel choices. There are and likely always will be severely limited resources with which to measure everything about available travel choices.

In summary, in deducing values from travel behavior, there is the dual problem of limited choices and limited measurements of the characteristics of the available choices. Both severely hamper the ability to deduce preferences and obtain the underlying values and trade-offs that govern travel behavior.

Deducing Values From Attitudes

Attitudes may be defined as people's "tendencies to respond in a particular manner to social or physical objects" (74), or their "disposition to react towards alternatives either positively or negatively" (70). In the sense of defining and filling in the gaps in the existing and potential output space within which travel decisions are made, information from the so-called transportation attitude survey literature can be helpful. Measurement problems with existing and contemplated surveys can be bypassed for immediate travel-research purposes, and the limitations of current modes can, to a limited extent, be ignored. However, the use of attitude survey results assumes that expressed attitudinal values can be related to behavioral and to corresponding measurement values.

On the relationship of attitudinal values to behavior, the most ambitious study to date, the Maryland study, found "a modest positive linkage between expressed attitudes and reported behavior, particularly for the work trip" (130). Wachs (176), in another study, reported,

> Drivers seem to be able to satisfy their preferences for many route characteristics. Drivers who express preferences for many route characteristics actually tend to travel on routes which possess them, whereas drivers who express little preference for such characteristics tend to drive on routes which do not possess them.

Watson (180) reported observing a positive relationship between estimated coefficients on modal-choice variables and the stated relative importance of those variables from a single attitudinal question asked together with those on behavior. Such positive general statements of the relationship are supported by the variations in reported attitudes of different mode- and route-user groups toward the modes and routes in question and toward their attributes (83).

On the relationship between attitudinal values and corresponding measurement values, the picture is not so clear. The results of attitude surveys must be used with caution. The Maryland study reported that "the importance of a particular attribute is a function of both the underlying strength of the human need or needs it is related to, and its present satisfaction level" (83). This function shows up in the results of many of the attitude studies in the literature. It requires that the use of attitude study results on the relative importance of attributes be weighted by current satisfaction levels. Otherwise, the importance of current, poorly satisfied modal attributes will be overestimated. This means that relative attribute importance ratings from one survey cannot be used as attribute weighting factors in a travel demand model separate from or independent of the values of the attributes in a proposed transportation system. Unfortunately the same can hold true for revealed preference (economic) models. We may safely agree with Wallace (177) that "one area which still needs considerable research because of the need for demand elasticity estimates is that of determining the relationship between attribute satisfaction ratings and the levels of attributes."

Problem 2: Choice Abstract Versus Choice Specific

The problem of choice abstract versus choice specific is a subset of the scaling problem. It is the question of the relationship of the stimuli to the travel choices: Are travel attributes perceived by themselves, or are they mapped on particular supply side choices such as mode and route or choice of technology? The question similarly can be extended to attributes of alternative destinations. These alternate perceptions of the travel environment imply that attributes of the transportation system can be included in travel-demand models in 1 of 2 ways: as choice-abstract or choice-specific attributes (21). For example, the gravity model as conventionally applied in transportation studies is a choice-abstract destination-choice model both with respect to its trip-attraction variables and its travel-time variables.

The significance of the question as a separate research focus in this project is that, if attributes are choice abstract, then the numbers of interaction terms in a travel model will be greatly reduced and the model will be much simpler to estimate and apply. Thus the possible payoff is great, and the question becomes well worth researching.

Lancaster (97) provided the classic statement of the choice-abstract concept in economics.

> Utility or preference orderings are assumed to rank collections of characteristics and only to rank collections of goods indirectly through the characteristics that they possess. . . .Furthermore, the same characteristic may be included among the joint outputs of many consumption activities so that goods which are apparently unrelated in certain of their characteristics may be related in others. (The traveler is assumed to derive utility, U, from the attributes, Z, consumed and obtained as a result of the transportation activity.)

A variation on the choice-abstract concept overlaps research problems 4 and 5, which relate to choice structure. This is the possibility that at least some attributes may be considered independently not only of 1 choice (such as mode) but also of all subchoices. The attributes may themselves provide the choice structure. For a simultaneous-choice model, the resulting choice-abstract model would be the same as a simultaneous model of choice-abstract attributes. However, if the choice-abstract attributes are themselves considered sequentially, a new choice structure results that corresponds to a recent model in mathematical psychology, the elimination-by-aspects model (173). The model is relatively easy to apply to travel choices (21), and, for this reason alone, problem 2 becomes significant and worth investigating. In addition, the elimination-by-aspects method may have its greatest usefulness in defining separable (strict-utility) models (problem 4) (21).

Problem 3: Measurement

Whether and how perceived values of attributes relate to physically measured values of attributes is the third scaling issue and has specific measurement implications. The question is not new. Transportation studies over the last 20 years confronted it, not very directly, when plotting perceived travel times from home-interview surveys against objective engineering estimates from skimmed trees. The correlation was observed to be considerably less than 1. Some data have been collected specifically for the purpose of relating perceived and engineering estimates of the times and costs of travel (100). However, the problem generally has been ignored until recently when data at the individual traveler level used to estimate (disaggregate) models have shown poor results. Ignoring the issue can result in biased parameter estimates with aggregate models and nonsignificant parameters with disaggregate models. The issue can be ignored no longer.

The problem of measurement is a classic problem in psychophysics, the branch of experimental psychology that relates stimulus to sensation or, in more conventional transportation terminology, the physical measurement of the stimulus to the perceived magnitude. The Stevens power law (159), which relates the magnitude of the sensation to the magnitude of the stimulus, may have important ramifications for the mathematical form in which physically measured variables should be included in travel models. It already has been applied to travel forecasting (with as yet uncertain results) by Ewing (52).

In environmental psychology and the newer subfield of psychogeography, a substantial body of literature relevant to travel has been accumulated (48). For example, it has been found that spatial judgements are influenced by traverse time and the quality of that time. Orme (138) found that "if two parts of a journey are of equal distance, that part will generally seem greater which is traversed at a slower speed for a longer time." Filled time and active participation in complex activity (driving, for example) make time seem to pass faster and possibly at considerably lower behavioral cost or disutility. The converse may be expected to be true for passengers in cars or transit vehicles, particularly if they are not permitted (or not inclined) to read or engage in some form of activity or stimulation. Recent work with children (73) shows that we learn about the environment by active manipulation of it (for example, by driving rather than by being driven), which indicates some profoundly different consequences on residential-location and car-purchase decisions of travel time by different available modes. (Thus there is an overlapping with research problem 2.)

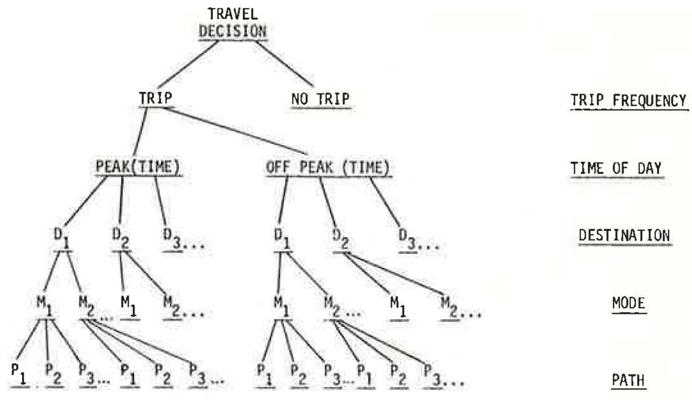
The measurement problem can be attributed, at least in part, to a lack of information on travel costs and opportunities as distance increases. Thus travel may decline with distance in the absence of any additional travel costs of distance. If information levels decrease with distance, then travel models may be falsely attributing declining travel interaction rates entirely to increased travel costs. Hanson (70) reports that "the set of known opportunities (the cognitive opportunity set) constitutes a rather limited proportion of the total opportunity set."

The measurement problem relates also to the question of how changes in the transportation system affect perceptions of the cost of movement. A better understanding of the measurement problem as an aspect of travel behavior could lead to an ability to organize transportation improvements in such a way as to transform the spatial pattern of travel demand in potentially more desirable ways for society.

Problems 4 and 5: Choice Structure

Problems 4 and 5 are the major choice-structure research problems. Are travel decisions decomposed into smaller decisions and, if so, how? Problem 4 relates to the

Figure 2. Sample presumed hierarchy of travel choices.



issue of separable choice, and problem 5 relates to the issue of possible independence among travel choices (and the mobility choices of problem 7). Quite different travel demand models can be derived analytically from or shown to be consistent with alternate behavioral assumptions on whether and how travel decisions are decomposed or partitioned into choice subsets between or within levels in the choice hierarchy shown in Figure 2 (21).

There are basically 3 levels of assumptions relating to the choice-structure issue. These are described in the context of the hierarchy of the short-run travel choices shown in Figure 2.

1. Simultaneous choice means that all attributes of the choice situation confronting the traveler are considered simultaneously. The complete trip is 1 decision. The relative valuation or relationships among the attributes are constant in any travel choice in the hierarchy shown in Figure 2. The hierarchy and the assumption could, of course, be extended to include longer term automobile-ownership and residential-location decisions.
2. In separability there is a set of travel decisions in which certain travel decisions are considered and may be modeled separately from other decisions. However, the relative valuation of choice attributes is constant in any complete travel decision (that is, any single path through the travel-decision tree shown in Figure 2).
3. In independence there is a set of travel decisions (choices) in which certain travel decisions are made independently of other decisions. Thus the relative valuation of choice attributes common to 2 or more travel choices is not likely to be the same in successive travel choices.

The first structure assumption is the simultaneous-choice assumption. The number of explanatory variables and the allowable interactions among variables that may be needed to explain (model) simultaneous travel behavior can multiply rapidly for realistic travel-choice situations in urban areas. The second and third assumptions restrict the attributes the traveler is assumed to evaluate in making his or her travel decision. Restricting the choices that are presumed available to the traveler is an appealing and popular way in which the complexity of travel demand models can be reduced. However, this involves making some important assumptions on the separability and possible independence of travel choices. The second assumption (separability) requires that the relative valuation of choice attributes be constant throughout the set of travel choices. Models of some subset of travel decisions can be calibrated based only on the subset of attributes describing those choices (33). The estimated utilities are then included in the models of the larger set of choices. Use is made of the separability property of strict-utility models derived from the independence of irrelevant alternatives axiom (110) under this recursive modeling method. Separability is a property (and therefore a behavioral assumption) of all share models in use today including the gravity model and the logit model (21).

There is some uncertainty and disagreement in the field on whether and when the assumption of sequential behavior is a requirement for the previously mentioned method of estimating or applying recursively estimated models. One school of thought holds that, when utilities are preserved in later estimated models of a larger set of travel choices, the sequence is assumed that is implied by the order of the conditional statement used to determine the choice attributes input to the utility function. That is, if, in Figure 2, utilities are estimated for the modal-choice decision by using modal travel costs that are defined by, or conditioned on, destination choice [that is, $P(M|D)$], and these utilities are preserved later in models of the larger set of choices, then sequential behavior is a required assumption. However, the opposite view is that the sequential estimation of utilities implies no such sequential behavior assumption. According to this second school of thought, the method of inclusive prices is simply a first approximation to the utility function estimated simultaneously over all the travel choices, and no sequential behavior is assumed (47).

There is, however, no disagreement on the requirement that separability must be assumed when one estimates share models on a subset of the entire set of travel choices or when one forecasts to a set of choices (such as destinations) that is larger or is a different partitioning from the choice set used in model estimation. Separability (the property that the relative probability of choice between 2 alternatives is independent of the presence of third or additional alternatives) constitutes a strong behavioral assumption by itself. Travel choices must be defined that may be considered substitutable for each other and not special cases of each other, or the separability assumption will be too strong. For example, the results of a gravity model forecast (which presently assumes destination choice as being separate and independent from the remaining choices) are vitally dependent on which destinations are considered to be substitutable for each other and which are considered to be special cases of each other. In current practice, all destinations are considered simple substitutes for each other. This is clearly contrary to the previously stated finding that "the set of known opportunities . . . constitutes a rather limited proportion of the total opportunity set" (70). Modal-split and path-choice models make the same complete substitutability and perfect information assumptions. Rules for logically restricting choice sets could considerably simplify the mechanics of travel forecasting as well as greatly improve the accuracy of travel forecasts.

The third assumption is the current assumption of urban transportation planning models that sequentially and independently estimate the different travel choices with different valuations of the independent variables in each model. The assumption does not by itself place restrictions on choice ordering. For example, the place of modal split in the order of trip-choice decisions has been called "the most actively debated issue in modal split" (181). However, application of the models in different sequences results in different sets of travel predictions. Also the assumption does not necessarily restrict independently modeled choices to levels (such as mode) in the choice hierarchy shown in Figure 2. They could be sets of alternative complete travel decisions. The third assumption may not be rejected on any sequence assumption so much as on the requirement that, for utilities derived from separately modeled travel decisions to be considered additive, their component attributes must be neither substitutes nor complements (have no interaction terms) (112). That is, if the utilities are assumed to be, strictly speaking, additive and independent sets of attributes governing each separate travel choice, the choices may indeed be modeled separately and independently. For example, some would say that automobile manufacturers appear to have successfully separated consumption of the automobile from consumption of travel. That is, at least until the current concern with energy-supply limitations came about, the attributes governing car-purchase decisions may have overlapped only insignificantly with attributes governing travel decisions.

In fact, if travel choices may be modeled separately only if their attributes are neither substitutes nor complements to attributes in the utility functions of other separately modeled choices, then the whole discussion of separately modeling the usually defined choices (Figure 2) perhaps should give way to structuring travel choices for modeling purposes by independent sets of choice-abstract attributes themselves (problem 2). If it can be demonstrated that travel choices are considered sequentially by travelers, the powerful evidence can be applied to support nonsimultaneous models of travel decisions. Three researchable hypotheses supporting sequential travel behavior have already been described (21):

1. Sequential choice ordering based on timing,
2. Sequential choice ordering based on adjustment time, and
3. Sequential choice ordering based on experience.

Other rationales for choice sequences no doubt also exist. There is considerable overlap here with the stochastic issue (problem 6) as well. That is, the results of research on problem 6 may provide considerable support for assuming certain kinds of choice structures under various circumstances.

The significance of the choice-structure issue is that considerable economies in time, money, and the use of existing skills and computer programs can be achieved if certain behavioral assumptions relating to travel-choice partitions can be shown to be acceptable under certain conditions. Also, to minimize some possibly grievous errors in application, the appropriate set of choices over which travel models should be applied in prediction is vitally dependent on an early resolution of this issue.

Problem 6: Learning

Problem 6 relates to the almost totally ignored possibility in transportation that travel patterns may be the manifestation of a series of conditioned behaviors and not simple behaviors stationary over time. Behaviorists, such as Skinner (152), would seriously question the integrity of anyone who thought otherwise. The basic truth of the "learning" hypothesis can hardly be questioned. Boulding (23) asserted that "a most serious defect" in transportation is to "assume simple preference or welfare functions on the one hand and opportunity functions on the other, without further inquiry and particularly without inquiring as to how these functions came into existence." Boulding accused most practitioners of believing in the "doctrine of the immaculate conception of the indifferent curve." He cited the well-established theory that,

> as people communicate with each other, individual preferences and value systems tend to converge into something which might almost be called a common value system. . . .Most people . . . are socialized into the society in which they grow up, accepting its preference structures and learning its technology.

Hartgen (74) reported that, "of all groups, the family is perhaps the most important in influencing individual behavior." Outside the household there are important differences in people's values, conditioned by the opportunities to which they have been exposed. On a very small scale, Wachs (176) found that "people's preferences for various route characteristics do vary, and the variations can be related to the characteristics of the people, their trips and the routes to which they have been exposed." On the urban scale, McMillan and Assael (123, 124) found that

> people in the five rail mass transit cities (in the U.S.) placed a higher value on public transportation compared to people in the rest of the country. They held (relatively) less favorable attitudes toward the automobile both as a mode of transportation and in relation to the satisfaction of specific transportation and personal needs in terms of its social role.

Similar examples of differences in values and relative attitudes toward public and private transportation between cities have been reported (71, 101). Because of learning and conditioning, therefore, potential limitations exist on the complete transferability of even a behavioral travel demand model from one population to another. The major issue of problem 6, therefore, is not so much whether travel can be considered to be a set of conditioned behaviors but what conditioning theory can tell us about changes in the transportation system and other circumstances surrounding travel that are likely to produce changes in travel behavior. Research on this issue, although it may be difficult (27), can be of great value in identifying attributes that influence travel choice (the scaling problem) and in yielding evidence on the structure of choice (as noted for problems 4 and 5). However, more specifically, research on this question can identify specific nonlinearities in the effects of certain attributes (threshold effects), variability or reliability from trip to trip in attributes (effect of reward schedule), and changes in direction or change in the circumstances surrounding the travel. With regard to the last, Shaffer (147) reported that, in the field of job and personnel research, it has been "demonstrated that job satisfaction is a function of the presence of variables different from those that are present in job dissatisfaction.... Correction of the dissatisfying elements is not enough to produce satisfaction."

Support for the hypothesis that improvement in the travel-cost dimension (dissatisfiers) achieves limited results in increasing mode usage was suggested by Rosinger et al. (144):

> It should be recognized that lists of negative attributes of transit do not necessarily imply appropriate positive actions. There may be better ways to generate demand than simply to concentrate on making transit 'unbad'. . . .There is no a priori reason why transit might not develop its own set of positive attributes in comparison with the auto. Certainly such an approach is more likely to generate patronage than is an absence of negative orientation.

Unfortunately, it is not clear that transit currently has any of its own set of positive attributes (such as attributes on which users score it higher than the automobile) (83). Attitudes and behavior such as this appear to be explainable by the concepts of conditioning theory.

There are several concepts from learning and conditioning theory that provide fruitful approaches to research on problem 6.

1. In stimulus generalization, the traveler may generalize conditioned travel behavior even in different circumstances. For example, the car that was used to "find" the residential location (active manipulation of space) continues to be used for the journey to work even though the utility of the trip by transit is greater than the utility by automobile.
2. In discrimination learning by contingencies, the traveler may learn to associate a different set of utilities with work trips than shopping trips and a different set with day trips than night trips.
3. In the reward schedule, the resistance to behavioral change by the traveler is a function of both quantity of rewards (the utility set consists of time, cost, comfort, and the like) and the frequency and pattern with which rewards are administered. In general, partial rewards are more difficult to extinguish than full rewards; for example, some bad trips will not extinguish certain behavior.
4. Satisfaction generally results in a conservative bias in the system of choice. That is, over time, levels of aspiration tend to adjust to levels of achievement. (The difference in levels is said to motivate search for new alternatives.) A new alternative may or may not change the traveler's perception of difference between current and possible (future) alternative states if he or she changes travel behavior. The reward schedule will have an important effect.

The 2 primary objectives of research on problem 6 are whether and how to incorporate learning into travel models and to discover what conditioning can tell us about travel behavior as it relates to all the other research questions (such as the formation of values governing travel-choice decisions in households, which is problem 7).

Problem 7: Decision Making by Large Behavioral Units

Problem 7 concerns how the circumstances of the household or other units affect travel behavior. Certain kinds of research questions relating to long-term household behavior and short-term travel behavior are similar. For example, just as short-term travel models may or may not assume that the marginal value of time and money is the same over the various travel choices (separability issue of problem 4), long-term models of household- and employment-location and car-purchase decisions may or may not assume the same marginal value of time and money for all categories of expenditures (including travel). And long-term choices may or may not be interdependent on each other or on short-term (travel) choices (problem 5) depending on the stimuli affecting those choices, their interactions, and their correlations.

It may be hypothesized that individual travel behavior results from the individual's perception of the quality of the transportation choices available to him or her that he or she evaluates in light of the requirements of a set of tasks or activities assigned by the household. It may be further postulated that these tasks are selected by the household through a decision process that considers the various needs of its members and the various alternative activities that may satisfy those needs and the constraints imposed by the cultural and financial circumstances of the household. The process may be said to have 3 components. First, through past experience, the individual selects alternative activities that satisfy his or her needs. Second, the activities and resources allocated to each household member are defined. Third, the elements of individual travel behavior are structured and performed to complete the assigned task.

Hartgen and Tanner (75) summed up the interplay of individuals and the groups to which they belong:

> Each individual has associated with him a set of needs defined by the roles he assumes in his interaction with other persons and groups. Through experience, individuals and groups develop both awareness of and attitudes toward alternate courses of action that may satisfy needs. Through awareness, a person or group recognizes the existence of those particular actions offering some potential for satisfying needs.

Various categories of needs (such as physical, social, and psychological needs) exist, and the circumstances of the household will affect the awareness of and attitudes toward alternative courses of action to satisfy those needs. For example, physical and nonphysical characteristics of individuals and groups will have an effect on the formulation and perception of their needs (25, 116) and on the information levels that they have on alternative trip-end opportunities (17).

It can be hypothesized that the family decision-making process evaluates the activities preferred by its members along 2 dimensions: the importance of the need that the activity is intended to fulfill and the resources and opportunities available for the activity. The task of child rearing may be completed within the physical confines of the household. However, other tasks such as income earning normally are accomplished in other locations, which gives rise to travel.

In the context of household needs and allocation of scarce resources (time and money) each individual must make subsequent decisions for completion of his or her assigned tasks. It is hypothesized that he or she reviews these activities and resources in relation to his or her own needs and experiences and determines those tasks that will best satisfy both the household's and his or her personal requirements. He or she then arranges these activities in a sequence for completion.

This social psychology stream of research already has been the subject of reviews as it relates to travel and car-purchase decisions (76). A second stream of research is the set of more purely economic investigations of household budgets, their changes over time, and their changes with respect to varying household circumstances, location in the metropolitan areas, and proximity to public transportation services. Such consumer budgeting studies at the household level, can be helpful in determining how changes in circumstances at the household level affect travel behavior.

A research approach that included household-level examination of problems 1 through 6 allows considerable perspective to be gained on the behavior of individual travelers. Households have sets of final demands and ways of accomplishing those final demands that are not available to individual travelers. The trade-offs within the household for time, money, and other resources are not visible and researchable at the individual-traveler level. Research on problem 7 can be expected to yield findings and insights that will be valuable both in terms of developing substantially new approaches to household-level travel modeling and in terms of clarifying issues and approaches to research on the scaling, structure, and learning questions.

CONCLUSIONS

Each of the 7 research problems is important in its own right, and all, of course, are interrelated. An improved understanding of travel behavior as articulated by each question can be a great help in efficient use of travel-forecasting models, and in appropriate selection of what forecasting model to use in the first place. Lack of knowledge of how travelers behave under different circumstances when confronted with high- versus low-capital-intensive transportation alternatives may be requiring the profession to always apply very complicated models. This has been called model overkill.

Four types of criteria are appropriate for evaluating alternative behavioral approaches to travel theory:

1. Travel behavior criteria ("truth");
2. Ability to incorporate behavior in a model;
3. Ability of the behavioral model to improve the accuracy of the conditional forecasts (policy and issue responsiveness); and
4. Planning process considerations (such as time, cost, data needs, and transparency).

The 4 types of criteria are divided into endogenous (item 1) and exogenous (items 2, 3, and 4) classes. Endogenous means evaluation on the basis of internally generated indicators. A travel demand model, like any model, is ultimately a subjective imitation of reality. Ultimately the modeler's understanding of behavior in the system of interest must be the starting point. Evaluating the "truth" endogenously means evaluating how well the theory incorporates the essential phenomena addressed in the first place. The 7 research problems represent 7 useful perspectives from which to evaluate endogenously the alternative behavioral approaches. It may be superfluous to indicate that each of the resulting 7 endogenous criteria incorporates the usual list of internal evaluation criteria such as adherence to theory, internal logic, reasonableness, conformance with prior knowledge, and the like. The remaining 3 types of criteria are essentially exogenous to the research questions. They are the wholly pragmatic considerations of the usefulness of incorporating alternative behaviors in models applied to relevant policy issues.

The second criterion (modeling ability) is important because only a limited number of tractable models that employ behaviorally different underlying assumptions for making operational alternative approaches to travel behavior appear to exist (<u>10</u>, <u>16</u>, <u>21</u>).

Three related issues are involved in the third criterion; they fall under the headings of accuracy, option responsiveness, and effects. Accuracy issues relate to the accuracy and the detail in space and time with which models implementing the behavioral assumptions can replicate and forecast travel. Option-responsiveness criteria are, of course, central to the concerns of travel forecasting and transportation planning. That is, for behavioral models to estimate travel and related consequences of transportation options, variables (attributes) describing those options must be included in the set of stimuli influencing travel behavior (the scaling problem). However, option sensitivity is only 1 criterion. If the option truly does not affect travel, the model should not be abandoned or distorted in some way to make it policy responsive. Knowledge of such travel-behavior insensitivity will have important consequences in policy analysis and evaluation. Effect-related criteria are a third type of policy-responsiveness criteria. It is important to clearly draw the distinction between travel forecasting and forecasting other effects. However, insofar as improved forecasts of travel can lead, for example, to improved air quality forecasts, the effect criteria are important for evaluating alternative behavioral approaches to travel theory.

The fourth criterion for evaluating alternative behavioral approaches to travel theory is planning process considerations. The practicing transportation planner is confronted with a bewildering proliferation of travel forecasting models and an increasing array of transportation planning options on whose consequences he or she must provide information. The growing number of options and the growing involvement of citizens in planning are resulting in greatly increased information requirements for decision making, shrinking time available for travel forecasting and forecasting other than strictly travel consequences, and greatly increasing "transparency" requirements placed on technical transportation planning procedures. Travel forecasting is not standing up well to any of these new developments. Although the second and third types of criteria act to a certain extent as constraints on what behaviors can be modeled and what policy issues can be accurately addressed, it is on the basis of this fourth criterion that the decisions in the field are going to be made concerning which travel forecasting techniques will be used in particular planning situations. An initial list of types of process criteria might include:

1. Time requirements,
2. Cost requirements,
3. Forecastability of independent variables,
4. Transparency of the models,
5. Hardware (computing) requirements,
6. Skill requirements,
7. Universality of the modeled behavior (such as transferability over time and space),
8. Reproducibility of the forecasts,
9. Provision of intermediate output, and
10. Ease of aggregating or disaggregating output.

It is clear that considerable research on travel behavior and the application of behavioral travel models will be required before specific planning trade-offs between the basis in behavior (and thus the logic and plausibility of travel forecasts) and the time, money, and skills required to carry out the forecasts can be made. However, such trade-offs are made either knowingly or unknowingly each time a travel-forecasting technique is applied. The research approaches suggested in this paper are directed toward making these trade-offs from a more informed basis.

ACKNOWLEDGMENT

I wish to sincerely acknowledge and thank Donald Olivier, Richard Brandon, Joseph Bain, and Gary Moore for their comments and assistance in previous work leading to the preparation of this paper.

REFERENCES

1. R. P. Abelson and M. J. Rosenberg. Symbolic Psychologic: A Model of Attitudinal Cognition. Behavioral Science, Vol. 3, 1958, pp. 1-13.

2. Qualitative Aspects of Urban Personal Travel Demand. Abt Associates, Inc., Cambridge, Mass., 1968.
3. R. Ackoff. Individual Preferences for Various Means of Transportation. Univ. of Pennsylvania, 1965.
4. T. J. Adler and J. A. Bottom. Formulation of Travel Demand Modelling Requirements. U.S. Department of Transportation, 1973.
5. E. Aldana. Towards Microanalytic Models of Urban Transportation Demand. Department of Civil Engineering, M.I.T., PhD dissertation, 1971.
6. W. Alonso. The Quality of Data and the Choice and Design of Predictive Models. HRB Special Rept. 97, 1968, pp. 178-192.
7. H. J. Bauer. A Case Study of a Demand-Responsive System. General Motors Corp., Warren, Mich., 1970.
8. M. Beckmann, R. L. Gustafson, and T. F. Golob. Locational Factors in Automobile Ownership. Annals of Regional Science, Dec. 1973, pp. 1-12.
9. M. Beckmann et al. Prediction of Demand for New Urban Transportation Systems. Proc., Urban Transportation Workshop, Operations Research Society of America, Baltimore, 1973.
10. M. E. Ben-Akiva. Structure of Travel Demand Models. M.I.T., PhD dissertation, 1973.
11. M. Ben-Akiva. Multidimensional Choice Models: Alternative Structures of Travel Demand Models. Engineering Foundation Conference, Berwick, Me., July 1973.
12. M. Ben-Akiva and M. Richards. Disaggregate and Simultaneous Travel Demand Models: A Dutch Case Study. Netherlands Ministry of Transport, 1974.
13. M. E. Ben-Akiva. Structure of Passenger Travel Demand Models. Transportation Research Record 526, 1974, pp. 26-41.
14. J. A. Benshoof. Characteristics of Drivers' Route Selection Behavior. Traffic Engineering and Control, Vol. 11, No. 12, April 1970.
15. F. C. Bock. Factors Influencing Modal Trip Assignment. NCHRP Rept. 57, 1968.
16. R. D. Bock and L. V. Jones. Measurement and Prediction of Judgment and Choice. Holden-Day, Inc., San Francisco, 1968.
17. S. Bowlby. Spatial Variation in Consumer's Information Levels. Transportation Center, Northwestern Univ., Household Travel Behavior Rept. 4, 1972.
18. D. Brand, B. Barber, and M. Jacobs. Technique for Relating Transportation Improvements and Urban Development Patterns. Highway Research Record 207, 1967, pp. 53-67.
19. D. Brand. Urban Travel Modes and Travel Attributes: Literature Review of Attitude Surveys. Department of City and Regional Planning, Harvard Univ., Jan. 1971.
20. D. Brand. Theory and Method in Land Use and Travel Forecasting. Highway Research Record 422, 1973, pp. 10-20.
21. D. Brand. Travel Demand Forecasting: Some Foundations and a Review. HRB Special Rept. 143, 1973, pp. 239-282.
22. D. Brand, E. Li, and D. Alshuler. A Disaggregate Direct Demand Model of Urban Work Travel. In The Interaction Between Urbanization and Land: Quality and Quantity in Environmental Planning and Design, Harvard Univ., Dec. 1973.
23. K. E. Boulding. The Formation of Values as a Process in Human Learning. HRB Special Rept. 105, 1969, pp. 31-49.
24. O. G. Brim et al. Personality and Decision Process. Stanford Univ. Press, Stanford, Calif., 1962.
25. R. Brown. Social Psychology. The Free Press, New York, 1965.
26. G. A. Brunner et al. User Determined Attributes of Ideal Transportation Systems. Department of Business and Management, Univ. of Maryland at College Park, 1966.
27. P. Burnett. The Dimensions of Alternatives in Spatial Choice Processes. In Geographical Analysis, pp. 181-204, July 1973.
28. P. Burnett. A Bernoulli Model of Destination Choice. Transportation Research Record 527, 1974, pp. 33-44.
29. A Behavioral Model of Automobile Ownership and Mode of Travel. Cambridge Systematics, Inc., Cambridge, Mass., 1974.

30. J. D. Carroll. Individual Differences and Multidimensional Scaling. In Multidimensional Scaling: Theory and Applications in the Behavioral Sciences (R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds.), Seminar Press, New York, Vol. 1, pp. 105-155.
31. J. D. Carroll and J. J. Chang. Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of "Eckart-Young" Decomposition. Psychometrika, Vol. 31, 1966, pp. 33-42.
32. F. S. Chapin and S. F. Weiss. Urban Growth Dynamics. John Wiley and Sons, New York, 1962, pp. 359-399.
33. A Disaggregate Behavioral Model of Urban Travel Demand. Charles River Associates, Inc., 1972.
34. A Model of Urban Passenger Travel Demand in the San Francisco Metropolitan Area. Charles River Associates, Inc., Cambridge, Mass., 1967.
35. J. S. Coleman. Models of Change and Response Uncertainty. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.
36. C. H. Coombs, R. M. Dawes, and A. Tversky. Mathematical Psychology: An Elementary Introduction. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1970.
37. J. E. Curtin. Effect of Fares on Transit Riding. Highway Research Record 213, 1968, pp. 8-20.
38. H. S. David. The Method of Paired Comparisons. Charles Griffin and Co., Ltd., London, 1963.
39. G. Debreu. Topological Methods in Cardinal Utility Theory. In Mathematical Methods in the Social Sciences (K. J. Arrow, S. Karlin, and P. Suppes, eds.), Stanford Univ. Press, Stanford, Calif., 1960.
40. T. B. Deen, W. L. Mertz, and N. A. Irwin. Application of a Modal Split Model to Travel Estimates for the Washington Area. Highway Research Record 38, 1963, pp. 97-123.
41. R. Dial. A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration. Transportation Research, Vol. 5, No. 2, 1971.
42. L. E. Haefner and L. V. Dickinson, Jr. Preliminary Analysis of Disaggregate Modeling in Route Diversion. Transportation Research Record 527, 1974, pp. 66-72.
43. L. B. Dinklage. Adolescent Choice and Decision-Making. Center for Research and Development on Educational Differences, Harvard Univ., 1967.
44. R. Dobson, T. F. Golob, and R. L. Gustafson. Multidimensional Scaling of Consumer Preferences for a Public Transportation System: An Application of Two Approaches. Socio-Economic Planning Sciences, Vol. 8, 1974, pp. 23-36.
45. T. A. Domencich, G. Kraft, and J.-P. Valette. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968, pp. 64-78.
46. T. A. Domencich and G. Kraft. Free Transit. Heath Lexington, Lexington, Mass., 1970.
47. T. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland Publishing Co., Amsterdam, 1975.
48. R. M. Downs and D. Stea. Image and Environment: Cognitive Mapping and Spatial Behavior. Aldine-Atherton, 1973.
49. O. D. Duncan and L. F. Schnore. Cultural, Behavioral and Ecological Perspectives in the Study of Social Organization. American Journal of Sociology, Vol. 65, Sept. 1955.
50. R. J. Ebert. Sequential Decision-Making: An Aggregate Scheduling Methodology. Psychometrika, 1971, pp. 303-316.
51. R. H. Ewing. Psychological Theory Applied to Mode Choice Prediction. Transportation, Vol. 2, 1973, pp. 391-410.
52. R. H. Ewing. The Nature of Travel Decision-Making. Transportation Research Record 526, 1974, pp. 66-72.
53. Urban Transportation Planning, General Information and Introduction to System 360. Federal Highway Administration, U.S. Department of Transportation, 1970.

54. H. Feger and B. Feger. Contributions to the Content Analytical Investigation of Decisions. Archiv für die Gesamte Psychologie, 1969, pp. 205-254.
55. M. J. Fertal et al. Modal Split. U.S. Bureau of Public Roads, 1966.
56. L. Festinger. A Theory of Cognitive Dissonance. Row, Peterson, New York, 1957.
57. N. N. Foote, ed. Household Decision-Making. New York Univ. Press, New York, Vol. 4, 1961.
58. C. R. Fleet and S. R. Robertson. Trip Generation in the Transportation Planning Process. Highway Research Record 240, 1968, pp. 11-31.
59. F. Goble. The Third Force: The Psychology of Abraham Maslow, A Revolutionary New View of Man. Pocket Books, New York, 1970.
60. R. G. Golledge and G. Rushton. Multi-Dimensional Scaling: Review and Geographic Applications. Commission on College Geography, Association of American Geographers, Washington, D.C., Technical Paper 10, 1972.
61. T. F. Golob and M. Beckmann. A Utility Model for Travel Forecasting. Transportation Science, Vol. 5, No. 1, 1971.
62. T. F. Golob et al. An Analysis of Consumer Preferences for a Public Transportation System. Transportation Research, Vol. 6, No. 1, March 1972, pp. 81-102.
63. T. F. Golob. Attitudinal Models. HRB Special Rept. 143, 1973, pp. 130-145.
64. T. F. Golob and R. Dobson. The Assessment of Preference and Perceptions Toward Attributes of Transportation Alternatives. Conference on Issues in Travel Demand Modeling and the Valuation of Travel Time, Berwick, Me., resource paper, July 1973.
65. P. E. Green. Marketing Research Applications of Nonmetric Scaling Methods. In Multidimensional Scaling: Theory and Applications in the Behavioral Sciences (R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds.), Seminar Press, New York, Vol. 2, pp. 185-210.
66. P. E. Green and R. F. Carmone. Multidimensional Scaling and Related Techniques in Marketing Analysis. Allyn and Bacon, Boston, 1970.
67. L. Guttman. A General Nonmetric Technique for Finding the Smallest Coordinate Space for a Configuration of Points. Psychometrika, Vol. 33, 1968, pp. 469-506.
68. D. Gurin. Travel Patterns and Problems of Suburban High School Males: Exploratory Study of a Population Subgroup With Recommendations. Harvard Univ., PhD dissertation, 1974.
69. D. G. Haney. Future Urban Transportation Systems: Desired Characteristics. Stanford Research Institute, Menlo Park, Calif., 1967.
70. S. Hanson. On Assessing Individuals' Attitudes Towards Potential Travel Destinations: A Research Strategy. Proc., Transportation Research Forum, 1974, pp. 363-370.
71. How Urban Americans View Their Transportation System. Louis Harris and Associates, Inc., Washington, D.C., 1969.
72. J. Harrison and P. Sarre. Personal Construct Theory in the Measurement of Environmental Images. Environment and Behavior, Vol. 3, 1971, pp. 351-374.
73. R. Hart. Children's Discovery in a Vermont Town. Clark Univ., PhD dissertation.
74. D. T. Hartgen. Mode Choice and Attitudes: A Literature Review. New York State Department of Transportation, 1970.
75. D. T. Hartgen and G. H. Tanner. Investigations of the Effect of Traveler Attitudes in a Model of Mode-Choice Behavior. Highway Research Record 369, 1971, pp. 1-14.
76. D. T. Hartgen. The Influence of Attitudinal and Situational Variables on Urban Mode Choice. New York State Department of Transportation, March 1973.
77. J. A. Hartigan. Direct Clustering of a Data Matrix. Journal, American Statistical Association, No. 67, 1972, pp. 123-129.
78. F. Heider. The Psychology of Interpersonal Relations. John Wiley and Sons, New York, 1958.
79. D. A. Hensher, P. B. McLeod, and J. K. Stanley. Comfort and Convenience: Theoretical, Conceptual and Empirical Extensions of Disaggregate Behavioral Models of Mode Choice. Melbourne, Australia, Feb. 1974.

80. L. R. Heusmann. A Study of Heuristic Learning Methods for Optimization Tasks Requiring Sequential Decisions. PhD dissertation, 1970.
81. Travel Behavior. Highway Research Record 446, 1973, 65 pp.
82. Urban Travel Demand Forecasting. HRB Special Rept. 143, 1973, 315 pp.
83. S. J. Hille and T. K. Martin. Consumer Preference in Transportation. Highway Research Record 197, 1967, pp. 36-43.
84. G. K. Ingram, J. F. Kain, and J. R. Ginn. The Detroit Prototype of the NBER Urban Simulation Model. National Bureau of Economic Research, New York, 1972.
85. N. Jardine and R. Sibson. Mathematical Taxonomy, John Wiley and Sons, New York, 1971.
86. J. F. Kain. A Contribution to the Urban Transportation Debate: An Econometric Model of Urban Residential and Travel Behavior. Review of Economics and Statistics, Vol. 46, 1964.
87. E. J. Kannel and K. W. Heathington. Temporal Stability of Trip Generation Relations. Highway Research Record 472, 1973, pp. 17-27.
88. E. J. Kannel and K. W. Heathington. Structural Model for Evaluating Urban Travel Relationships. Transportation Research Record 526, 1974, pp. 73-82.
89. H. Kassoff and D. S. Gendell. An Approach to Multiregional Urban Transportation Policy Planning. Highway Research Record 348, 1971. pp. 76-93.
90. M. Kemp. Some Evidence of Transit Demand Elasticities. Transportation, April 1973, pp. 35-51.
91. G. Kraft. Demand for Intercity Passenger Travel in the Washington-Boston Corridor. U.S. Bureau of Public Roads, Part 5, 1963.
92. G. Kraft and M. Wohl. Special Survey Paper: New Directions for Passenger Demand Analysis and Forecasting. Transportation Research, Vol. 1, No. 3, 1967.
93. G. Kraft. Alternative Travel Behavior Structures, Discussion. Transportation Research Record 526, 1974, pp. 52-55.
94. F. Koppelman. Prediction With Disaggregate Models: The Aggregation Issue. Transportation Systems Division, Department of Civil Engineering, M.I.T., working paper, 1974.
95. D. H. Krantz. Conjoint Measurement: The Luce-Tukey Axiomatization and Some Extensions. Journal of Mathematical Psychology, 1964.
96. J. B. Kruskal. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. Psychometrika, Vol. 29, 1964, pp. 1-27.
97. K. J. Lancaster. A New Approach to Consumer Theory. Journal of Political Economy, Vol. 74, 1966.
98. G. N. Lance and W. T. Williams. A General Theory of Classificatory Sorting Strategies: I. Hierarchical Systems. Computer Journal, Vol. 9, 1967, pp. 373-380.
99. G. N. Lance and W. T. Williams. A General Theory of Classificatory Sorting Strategies: II. Clustering Systems. Computer Journal, Vol. 10, 1968, pp. 271-277.
100. J. B. Lansing and G. Hendricks. Automobile Ownership and Residential Density. Institute for Social Research, Univ. of Michigan, 1967.
101. J. B. Lansing and G. Hendricks. How People Perceive the Cost of the Journey to Work. Highway Research Record 197, 1967, pp. 45-55.
102. J. B. Lansing, E. Mueller, and N. Barth. Residential Location and Urban Mobility. U.S. Bureau of Public Roads, June 1964.
103. W. Lassow. Effect of the Fare Increase of July 1966 on the Number of Passengers Carried on the New York City Transit System. Highway Research Record 213, 1968, pp. 1-7.
104. C. A. Lave. A Behavioral Approach to Modal Split Forecasting. Transportation Research, Vol. 3, No. 4, Dec. 1969, pp. 463-480.
105. H. LeBoulanger. Research Into the Urban Traveller's Behavior. Transportation Research, Vol. 5, 1971, pp. 113-125.
106. J. C. Lingoes. A General Survey of the Guttman-Lingoes Nonmetric Program Series. In Multidimensional Scaling: Theory and Applications in the Behavioral

Sciences (R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds.), Seminar Press, New York, Vol. 1, pp. 52-58.

107. P. S. Liou and A. P. Talvitie. Disaggregate Access Mode and Station Choice Models for Rail Trips. Transportation Research Record 526, 1974, pp. 42-52.

108. T. E. Lisco. The Value of Commuters' Travel Time: A Study in Urban Transportation Planning. Univ. of Chicago, PhD dissertation, 1967.

109. I. S. Lowrey. Seven Models of Urban Development: A Structural Comparison. HRB Special Rept. 97, 1968, pp. 121-145.

110. R. D. Luce. Individual Choice Behavior. John Wiley and Sons, New York, 1959.

111. R. D. Luce and H. Raiffa. Games and Decisions. John Wiley and Sons, New York, 1957.

112. R. D. Luce and P. Suppes. Preference, Utility and Subjective Probability. In Handbook of Mathematical Psychology (R. D. Luce, R. R. Bush, and E. Galanter, eds.), Vol. 3, John Wiley and Sons, New York, 1965.

113. M. Manheim. Fundamentals of Transportation Systems Analysis. M.I.T., preliminary edition, 1974.

114. M. L. Manheim. Practical Implications of Some Fundamental Properties of Travel-Demand Models. Highway Research Record 422, 1973, pp. 21-38.

115. J. G. March and H. A. Simon. Organizations. John Wiley and Sons, New York, 1958.

116. A. H. Maslow. Toward a Psychology of Being. Van Nostrand, Princeton, N.J., 2nd Ed., 1968.

117. J. P. Mayberry. Structural Requirements for Abstract Mode Models of Passenger Transportation. In The Demand for Travel: Theory and Measurement (R. Quandt, ed.), Heath, Lexington, Mass., 1970.

118. D. McFadden. Revealed Preferences of a Public Bureaucracy. Technical Institute of International Studies, Univ. of California, Berkeley, Technical Rept. 17, 1968.

119. D. McFadden and F. Reid. Aggregate Travel Demand Forecasting From Disaggregate Behavior Models. Department of Economics, Univ. of California, Berkeley, preprint, Nov. 1973.

120. J. M. McLynn et al. Analysis of a Market Split Model. Journal of Research of National Bureau of Standards, Vol. 72B, No. 1, 1968.

121. J. M. McLynn and T. Woronka. Passenger Demand and Modal Split Models. Northeast Corridor Transportation Project, U.S. Department of Transportation, 1969.

122. J. M. McLynn. The Metric Model. Office of High-Speed Ground Transportation, U.S. Department of Transportation, 1971.

123. R. K. McMillan and H. Assael. National Survey of Transportation Attitudes and Behavior: Phase I Summary Report. NCHRP Rept. 49, 1968.

124. R. K. McMillan and H. Assael. National Survey of Transportation Attitudes and Behavior: Phase II Analysis Report. NCHRP Rept. 82, 1969.

125. R. M. Michaels. Attitudes of Drivers Determine Choice Between Alternate Highways. Public Roads, Vol. 33, No. 11, Dec. 1965.

126. G. A. Miller, E. Gallanter, and K. H. Pribram. Plans and the Structure of Behavior. Henry Holt and Co., New York, 1960.

127. D. H. Mudarri. Mode Choice for Work Trips: A Key Urban Transportation Problem. Department of Economics, Univ. of Michigan, PhD dissertation, 1972.

128. H. A. Murray. Explorations in Personality. Oxford Univ. Press, New York, 1938.

129. A Survey to Determine Factors Which Influence the Public's Choice of Mode of Transportation. Joseph Napolitan Associates, Inc., Boston, Mass., Supplement 4 to Mass Transportation in Massachusetts, 1964.

130. A. N. Nash and S. J. Hille. Public Attitudes Toward Transport Modes: A Summary of Two Pilot Studies. Highway Research Record 233, 1968, pp. 33-46.

131. The Preference for Rapid Transit Among National Capital Region Commuters to Downtown D.C. In A Survey of Commuter Attitudes Toward Rapid Transit Systems, National Analysts, Inc., Philadelphia, Vol. 1, 1963.

132. Characteristics of Proposed New Rapid Transit Systems Which Will Maximize Riders. In A Survey of Commuter Attitudes Toward Rapid Transit Systems, National Analysts, Inc., Philadelphia, Vol. 2, 1963.
133. The Methodological Description of the Survey. In A Survey of Commuter Attitudes Toward Rapid Transit Systems, National Analysts, Inc., Philadelphia, Vol. 3B, 1963.
134. A. Newell, J. C. Shaw, and H. A. Simon. Elements of a Theory of Human Problem Solving. Psychological Review, Vol. 65, 1958, pp. 151-166.
135. G. C. Nicholaidis. Quantification of the Comfort Variable. General Motors Research Laboratories, Warren, Mich., preprint, 1974.
136. F. I. Nye and F. M. Bernardo. Emerging Conceptual Frameworks in Family Analysis. MacMillan Co., New York, 1966.
137. W. Y. Oi and P. W. Shuldiner. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, Ill., 1962.
138. J. E. Orme. Time, Experience and Behavior. Elsevier, New York, 1969.
139. F. T. Paine et al. Consumer Attitudes Toward Auto vs. Public Transport Alternatives. Journal of Applied Psychology, Vol. 53, No. 6, 1969, pp. 472-480.
140. J. Parker and J. M. Clark. Development and Validation of a New Model for Modal Choice. PTRC Urban Traffic Research Seminar, London, May 1972.
141. D. A. Quarmby. Choice of Travel Mode for the Journey to Work: Some Findings. Journal of Transport Economics and Policy, Vol. 1, No. 3, 1967.
142. R. T. Rao. Consumer's Purchase Decision Process: Stochastic Models. Journal of Marketing Research. 1969, pp. 321-329.
143. A. Rapoport. Optimal and Suboptimal Decisions in Perceptual Problems—Solving Tasks. Behavioral Science, 1969, pp. 453-466.
144. G. Rosinger et al. Design of Urban Transportation for the User. Battelle Memorial Institute, Columbus, Ohio, 1967.
145. F. Restle. Mathematical Models in Psychology. Penguin Books, Inc., Baltimore, Md., 1971.
146. S. Reichman and P. R. Stopher. Disaggregate Stochastic Models of Travel-Mode Choice. Highway Research Record 369, 1971, pp. 91-103.
147. M. T. Shaffer. Attitude Techniques in Action. Highway Research Record 305, 1970, pp. 112-122.
148. R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds. Multidimensional Scaling: Theory and Applications in the Behavioral Sciences. Seminar Press, New York, 1972.
149. A. Sherret. Structuring an Econometric Model of Modal Choice. Cornell Univ., PhD dissertation, 1971.
150. G. A. Shunk and R. J. Bouchard. An Application of Marginal Utility to Travel Mode Choice. Highway Research Record 322, 1970, pp. 30-39.
151. H. A. Simon. Models of Man. John Wiley and Sons, New York, 1957.
152. B. F. Skinner. Beyond Freedom and Dignity. Knopf, New York, 1971.
153. R. D. Smith. Heuristic Simulation of Psychological Decision Processes. Journal of Applied Psychology, 1968, pp. 325-330.
154. R. R. Sokal and P. H. Sneath. Principles of Numerical Taxonomy. W. H. Freeman and Co., San Francisco, 1963.
155. K. M. Solomon, R. J. Solomon, and J. S. Silien. Passenger Psychological Dynamics. In ASCE Sources of Information on Urban Transportation. Rept. 3, June 1968.
156. A. N. Sommers and F. F. Leikmuhler. A Nondemographic Factor V/STOL Prediction Model. Operations Research Society of America National Meeting, Philadelphia, Nov. 1968.
157. The Value of Time for Passenger Cars: Further Theory and Small-Scale Behavioral Studies. Stanford Research Institute, Menlo Park, Calif., 1970.
158. V. Stefflre. Some Applications of Multidimensional Scaling to Social Science Problems. In Multidimensional Scaling: Theory and Applications in the Behavioral Sciences (R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds.), Seminar Press, New York, Vol. 2, pp. 211-248.

159. S. S. Stevens. To Honor Fechner and Repeal His Law. Science, Vol. 133, 1961, pp. 80-86.
160. P. R. Stopher and T. E. Lisco. Modelling Travel Demand: A Disaggregate Behavioral Approach, Issues and Applications. Proc., Transportation Research Forum, 1970.
161. P. R. Stopher and A. Meyburg, eds. Issues in Behavioral Demand Modelling and the Valuation of Travel Time. Proc., Conference, Berwick, Me., Dec. 1973.
162. P. R. Stopher, B. Spear, and P. O. Sucher. Toward the Development of Measures of Convenience for Travel Modes. Transportation Research Record 527, 1974, pp. 16-32.
163. R. H. Strotz. The Empirical Implications of a Utility Tree. Econometrica, Vol. 25, No. 2, April 1957.
164. J. R. Stowers and E. L. Kanwit. The Use of Behavioral Surveys in Forecasting Transportation Requirements. Highway Research Record 106, 1966, pp. 44-51.
165. J. D. Tate and W. C. Howell. Term Expectation and Uncertainty in Human Decision Behavior. U.S. Air Force, USAF AMRL TDR No. 63-118.
166. H. Theil. Principles of Econometrics. John Wiley and Sons, New York, 1971.
167. T. C. Thomas and G. I. Thompson. Value of Time Saved by Trip Purpose. Highway Research Record 369, 1971, pp. 104-117.
168. L. L. Thurstone. A Law of Comparative Judgment. Psychological Review, Vol. 34, 1927, pp. 278-286.
169. W. S. Torgerson. Theory and Methods of Scaling. John Wiley and Sons, New York, 1958.
170. W. S. Torgerson. Multidimensional Scaling of Similarity. Psychometrika, Vol. 30, 1965, pp. 379-393.
171. R. C. Tryon. Cluster Analysis: (Correlation Profile and Orthometric Factor Analysis for the Isolation of Unities in Mind and Personality). Edwards Brothers, Ann Arbor, Mich., 1939.
172. L. R. Tucker and S. Messick. An Individual Difference Model for Multidimensional Scaling. Psychometrika, Vol. 28, 1963, pp. 333-367.
173. A. Tversky. Elimination by Aspects: A Theory of Choice. Psychological Review, Vol. 79, No. 4, 1972.
174. J. E. Vitt et al. Determining the Importance of User-Related Attributes for a Demand-Responsive Transportation System. Highway Research Record 318, 1970, pp. 50-65.
175. M. E. Vogel. Value of Travel Time: A Psychological Approach. Harvard Univ., Cambridge, term paper, 1973.
176. M. Wachs. Relationships Between Drivers' Attitudes Toward Alternate Routes and Driver and Route Characteristics. Highway Research Record 197, 1967, pp. 70-87.
177. J. P. Wallace III. Some Applications of Marketing Research Techniques to the New Mode Demand Forecasting Problem. Conference on Future Transportation Systems, Univ. of Pennsylvania, July 1969.
178. S. L. Warner. Stochastic Choice of Mode in Urban Travel: A Study of Binary Choice. Northwestern Univ. Press, Evanston, Ill., 1962.
179. P. L. Watson. Problems Associated With Time and Cost Data Used in Travel Choice Modeling and Valuation of Time. Highway Research Record 369, 1971, pp. 148-158.
180. P. L. Watson. The Value of Time and Behavioral Models of Modal Choice. Department of Economics, Univ. of Edinburgh, Scotland, PhD dissertation, 1972.
181. E. Weiner. Modal Split Revisited. Traffic Quarterly, Jan. 1969.
182. P. Wiers. Proposal for Travel Behavior Research to Professor Daniel Brand. Harvard Univ., Dec. 1972.
183. O. P. Williams. Metropolitan Political Analysis, A Social Access Approach. Free Press, New York, 1971.
184. H. J. Wooton and G. W. Pick. Trips Generated by Household. Journal of Transport Economics and Policy, Vol. 1, No. 2, 1967.

185. F. W. Young. TORSCA, an IBM Program for Nonmetric Multi-Dimensional Scaling. Journal of Marketing Research, Vol. 5, 1968, pp. 319-321.
186. F. W. Young. Polynomial Conjoint Analyses of Similarities: A Model for Constructing Polynomial Conjoint Measurement Algorithms. Psychometric Laboratory, Univ. of North Carolina, Rept. 74, April 1969.

DISAGGREGATE BEHAVIORAL MODEL OF AUTOMOBILE OWNERSHIP

Steven R. Lerman, Cambridge Systematics, Inc.; and
Moshe Ben-Akiva, Massachusetts Institute of Technology and Cambridge Systematics, Inc.

This paper describes a series of disaggregate behavioral models that forecast the probability that various combinations of automobile ownership and travel modes to work will be selected by households. The models assume that workplace and residential location are predetermined. The multinomial logit model is used in a joint structure that captures the complex interrelationship of automobile ownership and travel-to-work decisions. The paper describes the considerations in the choice of independent variables and the specification of the utility functions. The estimation results for each of 7 distinct socioeconomic groups, or market segments, with different behavioral characteristics are presented and analyzed. One of the models then is used to examine the shifts in automobile ownership for a suburban household resulting from alternative levels of transit service improvements.

FORECASTING automobile ownership has always played an important role in the transportation planning process. The car ownership of a household is a major determinant of its trip-making behavior. For example, car ownership has been found to be the single most important variable in trip-generation models (10, 12) and a major determinant of modal choice (3, 5, 7). Thus, to predict trip-making behavior adequately, one must first model the process underlying car-ownership decisions.

In most major transportation studies, car ownership typically has bridged the gap between urban land use models, which focus on the spatial pattern of urban activities, and the traditional 4-step transportation planning process of trip generation, trip distribution, modal split, and traffic assignment. This relationship between car ownership and the traditional transportation forecasting cycle is shown in Figure 1. Note that, in this figure, the transportation level of service should enter into the factors determining car ownership. However, despite its fundamental importance as a determinant of how people use an urban transportation system, forecasting car ownership often has been treated outside the basic forecasting cycle. It has been relegated to a side calculation made with simple models that rely on trend extrapolations or correlations between 1 and 2 variables and car ownership rather than on a strong causal theory. Few studies have ever addressed the basic behavioral factors underlying the household's car-ownership decisions, and still fewer have attempted to embody a behavioral perspective in a valid econometric model. Furthermore, no modeling efforts have attempted to incorporate in a behaviorally consistent way the interaction of the way in which households choose the number of cars they own and the other transportation-related decisions they make. This paper describes a study that developed a series of models that represent car-ownership behavior in a way consistent with a credible theory of the choice process at the household level. This series therefore can be used with reasonable confidence in a broad range of policy-testing situations.

The basic objectives of the study were

1. To formulate and test behavioral hypotheses about the relationships between car ownership and the transportation system,
2. To develop and estimate models of car ownership that can be used as part of an

overall transportation planning process and that are behaviorally structured and statistically valid,

3. To provide a tool for transportation policymakers to assess the impacts of policies on car ownership and in particular to assess the role that transit level of service plays in determining car ownership, and

4. To develop models that allow policymakers to separate the impacts of various transportation policies on different socioeconomic groups.

ISSUES IN MODEL DEVELOPMENT

Car ownership is a household-level decision that is highly dependent on how workers in that household choose to travel to work. Previous models (4, 6, 11) have failed to treat this interaction in a credible way; they have either ignored it entirely, assumed some arbitrary sequence of decisions, or used other highly simplistic behavioral assumptions.

Each household makes its car-ownership decision as part of a larger range of transportation-related choices. How many cars one owns depends on where one lives and how one uses those cars. For this reason, considering car ownership as part of a complex group of decisions that are hierarchically structured is desirable; that is, car-ownership decisions are made conditional on the outcome of some household choices and are made before the outcome of other decision processes. Figure 2 shows the choice hierarchy on which this study is based. The choices of employment location, residential location, and housing are assumed to be made on the longest term basis. Car ownership and travel mode to work, the medium-term choices, are made conditionally on the results of the long-term decisions. The short-term travel decisions of trip frequency, destination, mode, and the like, for various non-work-trip purposes are made last conditionally on the locational, work-trip, and car-ownership decisions.

This general hierarchy of choice has important implications for the way in which car ownership decisions should be modeled. Car ownership and mode to work are closely related and, therefore, to model them as jointly determined is most appropriate. (For this study, only the mode of the primary worker was considered. The primary worker, also termed the breadwinner, was defined as the working member of the household with the highest socioeconomic status on a scale developed during the study.) The overall model structure representing this choice hierarchy can be termed block conditional. This means that a block of lower-level choices is made conditionally on higher-level choices but that the choices within each block are determined jointly.

Another hypothesis on which the study was based is that households have very different underlying behavior depending on both the stage of life that they are in and the occupation of the primary worker. For this reason, different models are estimated for different market segments. The empirical results indicate that such market segmentation isolates groups of households that make car-ownership decisions by using very different sets of values and that previous modeling efforts that pooled all households may be biased substantially when used in forecasting.

In the initial model development, 9 distinct market segments were considered. These consist of 4 life cycle groups:

1. Households consisting of single persons without children,
2. Households with a married couple both of whom were younger than 45 years without children,
3. Households with children, and
4. Households with a married couple 1 or more of whom were older than 45 years without children.

Two occupational groups were defined as

1. Occupation A, which was made up of the blue-collar primary workers; and
2. Occupation B, which was made up of the white-collar primary workers.

These provided 8 market segments each labeled by life cycle and occupation. A ninth segment consisting of households with no full-time workers also was defined and labeled life cycle 5. In addition, a random sample of households in life cycles 1 through 4 was used in model estimation for comparison with the separate market-segment models.

The models described in this paper use disaggregate data from the 1968 Washington, D.C., home-interview survey to estimate a series of multinomial logit models. Choice theory and the logit model are not reviewed in this paper. Excellent discussions of the theory and use of disaggregate choice models are available elsewhere (2, 3, 8, 9).

The models developed in this study predict the probability with which a household will select each of a number of available alternative car-ownership and travel-mode-to-work combinations. Thus, by the multinomial logit form, the probability of a household's choosing to own a given number of cars a and traveling to work by mode m is

$$P(a, m) = \frac{\exp(V_{am})}{\sum\limits_{a}\sum\limits_{m} \exp(V_{am})} \tag{1}$$

where V_{am} is the utility of a car-ownership and travel-mode-to-work combination.

Given the joint model for the probability P(a, m), one can derive any desired conditional or marginal probabilities. For example, a simple car-ownership model with indeterminate modal choice will be

$$P(a) = \sum_{m} P(a, m) \tag{2}$$

and a model of modal choice given a car-ownership level will be

$$P(m \mid a) = \frac{P(a, m)}{P(a)} \tag{3}$$

The utility functions V_{am} are restricted to be linear in the coefficients.

CHOICE OF VARIABLES

There are 6 basic categories of variables that influence a household's car-ownership and travel-mode-to-work decision:

1. Transportation level of service to work,
2. Car-ownership costs,
3. Locational attributes,
4. Housing attributes,
5. Spatial opportunity variables, and
6. Socioeconomic variables.

Each of these categories represents important factors that determine car ownership and travel mode to work. Therefore, none should be ignored completely. However, the level of detail with which each category can be treated can vary widely. For example, housing can be treated as a 0-1 dummy variable to represent whether the household lives in a single-family dwelling, or it can be treated as a broad range of variables describing rent, type of tenure, years at residence, type of structure, age of structure, and type of lot. The trade-off between level of detail and relative ease of use in forecasting was basic to the problem of variable selection, particularly when one of the objectives of the study was to capture as much of the behavioral process as possible.

Note that not all the variables appear directly in the utility functions. Some variables are combined in ways that reflect a theory about how they interact. Others enter into the utility functions of some car-ownership and travel-mode-to-work pairs and not others. Thus this section is not intended to be a discussion of the actual independent variables used for estimation, but rather it is a list of those variables that in some form affect the probability of a household's selecting any alternative. What follows is a discussion of each category of variables and the actual measures used in the models to represent it.

Transportation Level of Service to Work

The variables for transportation level of service to work influence directly the primary-worker's choice of travel mode to work, and therefore influence car ownership through the interdependency of car ownership and travel mode to work represented by the joint structure of the model. It was decided to use the traditional level-of-service measures commonly included in modal-choice models: in-vehicle time, out-of-vehicle time, and out-of-pocket cost. Empirical tests indicated that the time spent out of vehicle apparently is perceived to be far more onerous for short trips than for longer ones. Hence the ratio of out-of-vehicle time to highway travel distance was used. The level-of-service measures of both the car and the transit mode appear in the model in their respective utility functions.

Even when these 3 types of level-of-service variables are used, certain effects that influence modal choice, such as the frustration of being in congested traffic and the high variance of peak-hour travel time in heavily used travel corridors, are not measured. It is these factors that typically cause modal choice for downtown-oriented trips to be different from non-central-business-district (non-CBD) trips (13). To capture this effect, we used a dummy variable for CBD-oriented work trips.

Car-Ownership Costs

Car-ownership costs are the primary deterrent to high car ownership for households with more than 1 driver. However, the cost of owning a car is highly dependent on the type of car owned. Small cars with their low purchase prices and good fuel economy are far cheaper to own on an annualized basis than are larger, less efficient cars. However, the focus of this study is on the total number of cars a household will own rather than on their type, age, or quality. Thus, some value representing the average cost of owning a specific number of cars must be used.

An average figure of $1,000 ownership cost/car/year was used in the models to be described under socioeconomic variables. However, because this figure was selected somewhat arbitrarily, tests were made to determine the sensitivity of the results to large changes in that cost. These results indicate that large variations in this assumed value do not substantially alter the parameter estimates.

Locational Attributes

The way in which locational attributes affect the car-ownership and travel-mode-to-work decision is far more subtle than the effects of the first 2 types of variables discussed. At the simplest level, differences in personal property taxes and insurance will add to the monetary cost of owning cars; however, these variations in the Washington, D.C., area were deemed too small to be relevant. What is more significant is that implicit in a household's location decision are important restrictions on the way in which households may travel to work or to shop. For this reason, no specific variables were introduced into the utility functions to represent locational effects; instead, location was reflected in the other variables, such as those for spatial opportunity and level of service to work, and in restrictions on the alternatives available to each household as described under spatial opportunities.

Housing Attributes

Housing attributes can be represented by a large number of possible variables. However, on careful consideration of the way in which a household selects the number of cars it wishes to own, it was decided that only 1 major influence, whether the household resided in a single-family dwelling, merited consideration. Single-family houses typically have driveways and generally are located in areas that have readily available on-street parking; multiple-family dwellings generally are characterized by the reverse. Thus one would anticipate that, if all else is equal, the utility of multiple-car ownership would be substantially higher for those residing in single-family homes than for those residing in apartments.

Spatial Opportunities

Spatial opportunities make up perhaps the most difficult class of variables to represent and measure. These variables are a composite of the attributes of all nonwork trips, and some way of combining the characteristics of various possible trips with the relative likelihood of the household's making them must be found. The approach used in this study relies on a behavioral means of combining the level of service to different nonwork destinations. In the choice hierarchy shown in Figure 2, when a household selects its car ownership and mode to work, its actual pattern of nonwork travel is indeterminate. However, given any choice of car ownership and travel mode to work, the household then would be able to determine the probability with which it would travel to each destination. These probabilities depend on the transportation level of service and the attractiveness of each destination, but they also depend on the characteristics of the household. Therefore, it makes theoretical sense to use an estimate of the household-level probabilities to weight the level of service for nonwork travel. Expressed mathematically, this spatial opportunity measure is determined as follows [a more detailed discussion of this type of composite variable is available elsewhere (2, 3)]:

$$\text{Accessibility of a household in zone i by mode m} = \sum_{d} t_{idm} \cdot P(i, d, m) \qquad (4)$$

where

$P(i, d, m)$ = probability of the household's traveling from i to d by mode m, and
t_{idm} = some measure of the level of service from i to d by mode m.

Actually measuring this type of accessibility gives rise to a number of practical issues such as what types of spatial opportunities should be used. It was decided that the most relevant nonwork travel purpose is shopping and that other spatial opportunities play a secondary role and reasonably can be ignored. This greatly reduced the computational problems of creating the accessibility measures without substantially sacrificing important car-ownership effects. Another practical issue concerns what measure of level of service is most appropriate. Although each travel-time and cost measure can be used separately, this gives rise to an unwieldy number of variables. Hence generalized prices that are a weighted linear function of in-vehicle time, out-of-vehicle time, and out-of-pocket cost were created. Furthermore, the value of time used in the weighting process was itself a function of income, reflecting the hypothesis that high-income households would be willing to pay more to save time than low-income households would. Because all the level-of-service measures are for shopping trips, which typically are made in the off-peak hours, they are substantially different from the work-trip level-of-service measures.

The last practical issue that arises from the use of this type of accessibility measure is how the probabilities P(i, d, m) and the parameters of the generalized prices can be estimated. Because the underlying motivation for using this type of variable is its behavioral, disaggregate interpretation, it seems logical that a previously estimated disaggregate choice model should be used. This is precisely what was done in this study. By using a model developed by Ben-Akiva (2), we determined based on off-peak level-of-service data the probabilities of every household in the sample to select each shopping destination. Ben-Akiva's model (2) is in itself a simultaneous-choice model; in this case, choice of mode and destination for shopping trips are considered. By using the joint probability of mode and destination, P(d, m), we derived the conditional probabilities P(d | car) and P(d | transit). The weighting parameters of the generalized prices were taken from the utility function of this joint probability model. By using the forecasting probabilities, we determined the attributes for the expected shopping trip by both car and transit. These values are a function of both the home zone and the income of the household and hence are highly disaggregate measures of spatial opportunities.

Socioeconomic Variables

Socioeconomic variables are a special class of attributes in that they do not vary across car-ownership and travel-mode-to-work alternatives. For this reason, these variables must somehow be transformed in the utility function either by combining them with other variables or by making them alternative specific (including them in the utility function of some car-ownership and travel-mode-to-work combinations and not in others). These transformations of the actual measures will be considered in the next section of this paper.

Household income clearly plays an important role in determining automobile ownership. Simple tabulations of automobile ownership for different income groups indicate a strong and positive correlation. Higher income would increase the relative utility of the more expensive alternatives such as owning more than 1 car and taking a car to work.

Household size also should play a role in the car-ownership decision. Large households typically will require a greater portion of their available income for essentials such as food, housing, and clothing, thus leaving fewer family resources for expenditures on automobiles.

The number of licensed drivers represents the competition among household members for the use of cars. The greater the competition is, the more likely it should be that the household will own a greater number of automobiles and the more likely it should be for the primary worker to use transit to work. Furthermore, a household is extremely unlikely to own more cars than it has drivers in a household; this places an upper bound on car ownership. Table 1 gives a summary of the variable categories that are introduced into the joint car-ownership and modal-choice model.

SPECIFICATION OF THE JOINT UTILITY FUNCTIONS

The previous sections consider only the variables included in the model. This section will address the equally important question of how these variables are represented in the utility functions of each car-ownership and travel-mode-to-work combination. All of the models discussed in this section assume that a maximum of 5 alternatives are available to any household:

1. 0 cars owned and transit taken to work,

2. 1 car owned and car taken to work,
3. 1 car owned and transit taken to work,
4. 2 or more cars owned and car taken to work, and
5. 2 or more cars owned and transit taken to work.

For convenience, the alternatives in items 4 and 5 will be referred to as having 2 cars even though they consist of 2 or more cars. Thus 5 utility functions (1 for each possible alternative) are to be considered. However, not every household in the sample has every alternative available.

Even when the restriction that the utility function must be linear in its parameters is imposed, the number of ways in which variables can be formulated is virtually limitless. Variables can be multiplied, added, or divided, or their logarithms can be used. Some variables may appear in one utility function and not in others. Furthermore, socioeconomic variables, because their values do not vary across alternatives, must be transformed somehow either by combining them with other variables or by making them alternative specific.

The first group of variables selected represents a constant term added to each utility function. A different constant term can be introduced into all but 1 of the utilities. These constants measure pure alternative effects, that is, the attributes of the alternative relative to the one without a constant term that are not measured in all the other variables. A constant term was introduced into each utility except for the case of cars owned and transit taken to work. This choice of which utility should not have a constant term is completely arbitrary and has no effect on the probabilities of selecting each alternative.

The next variable used reflects the fact that the level of car availability a household would have if it chose any particular alternative affects its perception of the desirability of the alternative modes to work. The number of cars per licensed driver pertaining to each alternative was used to measure this effect. Because this variable was selected to measure the modal-choice aspect of the decision process, it was defined as follows:

$$\text{Cars per licensed driver} = \begin{cases} \dfrac{\text{number of cars in the alternative}}{\text{number of licensed drivers in the household}} & \text{for car-to-work alternative} \\ 0 & \text{otherwise} \end{cases}$$

A variable of this type frequently appears in simple modal-choice models because of the presumed effect of car ownership on whether to take car or transit for a trip. As it is defined here, it plays precisely the same role, except that it now affects both modal choice (directly) and car ownership (through the simultaneity in the model structure). In addition, in modal-choice models, this variable reflects the chosen car ownership rather than the various levels that might have been selected. Thus an increase in the number of licensed drivers will cause a decrease in the value of the variable. One would expect that this should decrease the probability that the household will have its primary worker take the car mode to work; therefore, the coefficient of this variable should be positive.

The next variable requires some explanation. It arises from the fact that a large number of monetary measures are in the model, including household income, car-ownership costs, and out-of-pocket travel costs for the work trip. Furthermore, it was hypothesized that the way household size affects car ownership is by altering the amount of gross income available for nonessential expenditures. Clearly, one would like to avoid introducing a separate variable for each of these monetary factors. The question is how these attributes can be combined into a single variable representing the money that would be available to the household if it selected each alternative. This was done by formulating a variable, termed for reference as the remaining income variable, as follows:

Remaining income = household annual income - 800 (household size)
- 1,000 (number of cars in the alternative)
- 250 (out-of-pocket cost of the work trip of the alternative)

Thus the value of this variable is an approximate measure of the amount of money a household has left over after expenditures on essential goods [assumed to be $800/household member/year based on work by Mudarri (11) and Hoxie (6)], car-ownership costs ($1,000/car owned), and work-trip cost (250 round-trip work trips/year). An alternative with 0-car ownership results in a high value of remaining income, representing the availability of income the household otherwise would have to allocate to the purchase, maintenance, and operation of a car if it had chosen to do so. The coefficient of this variable in the utility function always should be positive to reflect the fact that, if all else is equal, households would rather have more money than less money. The use of the remaining income variable is a classic case of the application of variable selection criteria based on deductive reasoning. Its justification is based on the theoretical considerations rather than empirical ones. However, the motivation for developing this variable in the first place was that reliable estimates of the coefficients of the separate cost components could not be obtained, and excellent estimates of the remaining income variable coefficient could. At first glance, this would seem to be contradictory. However, by collapsing all costs into a single variable, we have added an extra piece of information to the model formulation; the marginal utility of any cost component has been assumed to be the same, regardless of the type of expenditure considered. Stated more simply, the household has been assumed to view a dollar as having the same value regardless of where it is spent. It is this assumption that improved the estimate of the cost-term coefficient. To do this, a piece of deductive information has been used.

It was decided that this variable should not enter the utility functions linearly; the utility a poor family derives from an extra dollar is probably much greater than that which a wealthy family derives. Thus the marginal utility of money should decrease as the value of remaining income increases. This hypothesis was reflected by using the natural log of remaining income as an independent variable rather than by using simply the value itself.

The next variable is the type-of-housing dummy variable. This was defined to be equal to 1 only in the 2-car alternatives for households residing in single-family dwellings. Thus the coefficient of the single-family-housing dummy represents an added utility to the multiple-car options for a specific group of households and leaves the utility of the remaining options unaffected. Presumably, the coefficient of this dummy variable should be positive.

The next variable is the in-vehicle travel time in minutes for the round-trip work trip. This variable reflects the disutility of travel time, and hence should have a negative coefficient. Note that this variable has the same value for all alternatives with the same mode to work for the primary worker and is independent of the car ownership associated with the alternative.

The other time variable, out-of-vehicle time, is measured with respect to the distance of the work trip. This variable is defined as follows:

$$\text{Excess time/distance ratio} = \frac{\text{out-of-vehicle time for round-trip work trip in minutes}}{\text{1-way travel distance in miles (kilometers)}}$$

This choice of an out-of-vehicle time measure was made largely on empirically based criteria. Other forms of the variable typically produced results that were less satisfactory, and no strong theoretical reason existed for selecting one over the other. Travel distance was measured along the highway network. As with in-vehicle time, the value of this variable is invariant among car-ownership levels.

The next variable was designed to reflect another effect of the number of licensed drivers within a household. Although number of licensed drivers affects choice of mode to work through the cars-per-licensed-driver variable, it also should affect the level of car ownership directly. The more licensed drivers there are in a household, the more likely it should be to select a high car-ownership level independently of the travel mode to work taken by the primary worker. This effect was measured by introducing a variable that reflects the number of licensed drivers into each utility function with a different coefficient for each car-ownership level. These variables were defined for 1-car alternatives as follows:

$$\text{Inverse licensed drivers}_1 = \begin{cases} \dfrac{1}{\text{number of licensed drivers in the household}} & \text{for 1-car alternatives} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Inverse licensed drivers}_2 = \begin{cases} \dfrac{1}{\text{number of licensed drivers in the household}} & \text{for 2-car alternatives} \\ 0 & \text{otherwise} \end{cases}$$

When these variables originally were introduced into the model, it was hypothesized that the effect for the 2-car alternatives (as measured by the coefficient value) would be twice as great as the effect for the 1-car alternatives. Statistical tests indicated that this was indeed the case, and, for the models ultimately selected, the 2 licensed driver variables were combined into a single variable, defined as follows:

$$\text{Inverse licensed drivers} = \begin{cases} 0 \text{ for the 0-car and transit-to-work alternative} \\ \text{inverse licensed drivers}_1 \text{ for the 1-car alternatives} \\ 2\,(\text{inverse licensed drivers}_1) \text{ for the 2-car alternatives} \end{cases}$$

The use of the inverse of the number of drivers rather than simply the number of drivers reflects the hypothesis that, as the number of drivers increases, the marginal effect of an additional driver on the need for automobiles decreases. Clearly, the coefficient of the inverse licensed drivers variable should be less than 0.

The next 2 variables represent the spatial opportunities for nonwork travel. In the household's selection of car ownership, the absolute level of shopping accessibility is of little importance. What is actually relevant is whether the household will typically use car or transit for shopping trips; therefore, the cost of using a car relative to the cost of using transit influences car ownership. For this reason, the following variable was developed.

$$\text{Generalized price ratio} = \frac{\text{expected generalized car cost for shopping travel}}{\text{expected generalized transit cost for shopping travel}}$$

As defined, this variable does not change value for different alternatives and therefore must be introduced into the utility function as alternative specific. (This would not have been true had the shopping-trip model used car ownership as an explicit dependent variable. However, it still might have been desirable to capture differences in the way accessibility is perceived for different car-ownership levels.) Thus the following 2 variables appear in the model.

$$\text{Generalized price ratio}_1 = \begin{cases} \text{generalized price ratio for 1-car alternatives} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{Generalized price ratio}_2 = \begin{cases} \text{generalized price ratio for 2-car alternatives} \\ 0 \text{ otherwise} \end{cases}$$

Note that the cost of taking transit when transit is not available (as in many suburban zones) is for practical purposes infinite; in such cases, the value of the generalized price ratio is 0. As generalized shopping travel cost by car increases, the value of the generalized price ratio increases. One would anticipate that this increase in car cost would result in greater use of transit. Consequently, the likelihood of high car ownership should decrease. To reflect this hypothesis, the coefficient of both these variables should be negative because they both measure the effect of shopping accessibility relative to the 0-car and transit-to-work alternative. Furthermore, the effect should be greater for the 2-car alternatives than for the 1-car options. This should result in a larger coefficient for the 2-car alternative than for the 1-car alternative.

The last variable used in the models is the work-trip-destination dummy variable, which is used to reflect the added disutility of traveling downtown by car not entirely captured by the level-of-service measures. This variable is defined as follows:

$$\text{CBD work-place dummy} = \begin{cases} 1 \text{ if work place is in CBD for car-to-work alternatives} \\ 0 \text{ otherwise} \end{cases}$$

Note that this variable is defined as specific to the car mode. Therefore, the coefficient reflects the difference in utility between car and transit for downtown work trips if all else is equal, and it should be negative.

Table 2 gives a summary of the variables and their definitions. However, the precise structure of each utility function is not clear and can best be illustrated by writing out the utility functions with the variables that are always 0 eliminated. For example, if the coefficients of the 13 variables are denoted B_1, B_2, and so on, the utility of variable 2 is as follows:

$$V_2 = \beta_2 + \beta_6 + \beta_8 + \beta_9 + \beta_{10} + \beta_{11}$$

The remaining 7 coefficients do not appear in the utility because the value of the variable with which they are multiplied is by definition 0. A more complicated utility function is that of variable 3 and is given in the following equation:

$$V_3 = \beta_3 + \beta_5 + \beta_6 + \beta_7 + \beta_8 + \beta_9 + \beta_{10} + \beta_{12} + \beta_{13}$$

Figure 1. Relationship of car-ownership model to conventional transportation forecasting cycle.

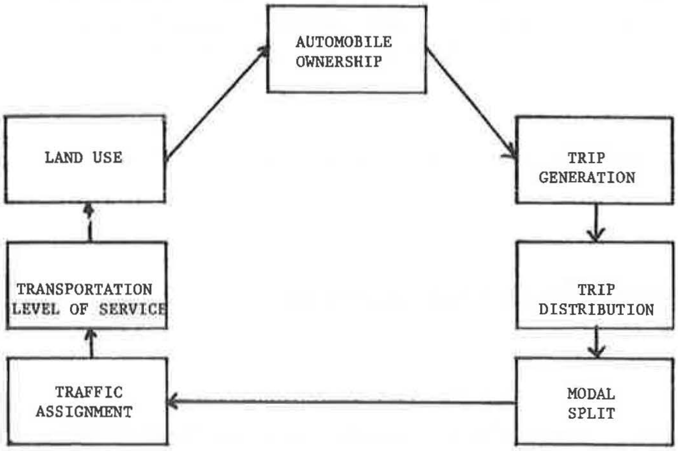


Figure 2. Choice hierarchy.

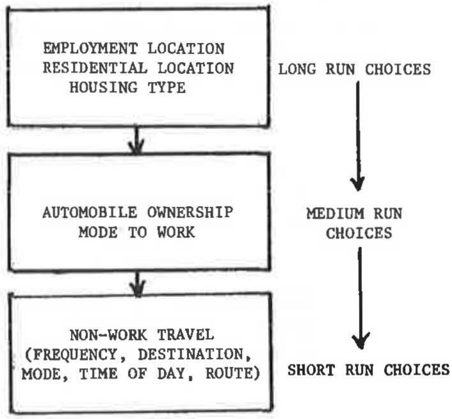


Table 1. Variables used in the model.

Category	Variables Used in Model
Level of service to work	In-vehicle time, out-of-vehicle time, and out-of-pocket cost by both car and transit Dummy variable representing downtown destination
Car-ownership costs	Cost of $1,000/car assumed
Locational attributes	Not explicitly introduced into model (level of service, traveling distances, and spatial opportunities are for residential and job locations)
Housing attributes	Dummy variable indicating single-family dwelling
Spatial opportunities	Expected generalized prices for shopping trips by both car and transit
Socioeconomic variables	Income, household size, and number of licensed drivers (life cycle and occupation are used for market segmentation)

Table 2. Variables and constants denoted by coefficient.

No.	Variable	Definition
1	1 car and car taken to work (constant)	1 for 1 car and car taken to work; 0 otherwise
2	1 car and transit taken to work (constant)	1 for 1 car and transit taken to work; 0 otherwise
3	2 cars and car taken to work (constant)	1 for 2 cars and car taken to work; 0 otherwise
4	2 cars and transit taken to work (constant)	1 for 2 cars and transit taken to work; 0 otherwise
5	Cars per licensed driver	Number of cars per licensed driver for car taken to work; 0 otherwise
6	ln (remaining income)	Household annual income - (800 × number of persons in household) - (1,000 × number of cars) - cost of 250 daily round trips
7	Single-family dwelling dummy	1 if household lives in single-family dwelling for 2 cars; 0 otherwise
8	In-vehicle time	Daily round-trip in-vehicle commuting time in minutes
9	Excess time/distance ratio	Daily round-trip out-of-vehicle commuting time in minutes/1-way distance of work trip in miles (kilometers)
10	Inverse licensed drivers	Number of cars/number of licensed drivers
11	Generalized price ratio$_1$	Generalized travel cost of shopping by car/generalized travel cost of shopping by transit for 1 car; 0 otherwise
12	Generalized price ratio$_2$	Generalized travel cost of shopping by car/generalized travel cost of shopping by transit for 2 cars; 0 otherwise
13	CBD workplace dummy	1 if work trip is CBD oriented for car taken to work; 0 otherwise

Table 3 gives a summary of the structure of all the utility functions. Each type of variable is given, and the coefficient with which this measure is multiplied appears for each alternative. Zeros appear where the variable has no effect on the utility of the alternative.

SET OF AVAILABLE ALTERNATIVES

The underlying choice theory of the logit model requires that the choice set consist only of feasible alternatives. This implies that, to properly estimate a joint car-ownership and modal-choice model, one must know which of the possible set of 5 alternatives is actually available to every household. However, data on the available alternatives were not included in the survey, and even if data had been included, they would have limited usefulness because of various reporting biases (14). This does not mean that the issue of available alternatives can be ignored. Even without data on reported choice set, it is possible to state with fairly high reliability that some households will not consider some alternatives as being available. For example, households without drivers only have the 0-car-ownership and transit-to-work alternative. Thus, no model need be estimated for them at all because they will always select the 0-car option with probability 1. A more interesting possible restriction of available options is that households living in fringe suburban areas do not have a transit-to-work option. These households would have only 2 alternatives: 1 car and car to work and 2 or more cars and car to work.

By using a series of rules such as these, one can approximate the set of feasible alternatives fairly well. This process has been termed screening the alternative set, and is an important part of the modeling process. Failure to do this will result in estimates that are biased and inconsistent, and, therefore, will result in unreliable forecasts of future conditions.

ESTIMATION RESULTS

By using the variables previously described, we estimated a number of joint car-ownership and travel-mode-to-work models. The first models developed were for each of the market segments defined in the section on issues in model development. However, some of these segments were too small for reliable results to be obtained. For this reason, segments 2A and 2B, blue-collar and white-collar young married couples without children, and segments 4A and 4B, blue-collar and white-collar older married couples without children, were collapsed into 2 segments corresponding to life cycles 2 and 4. Furthermore, segment 3B, households with children, was so large that, to reduce computational requirements, only half of the available data were used.

Some of the market-segment models are slightly different from the general form described in the section on specification of the joint utility functions. For example, life cycle 1 consists of single-person households; therefore, the 2-car alternative is not relevant. In the model for this segment, the single-family housing variable was redefined to apply to 1-car alternatives rather than the multiple-car options. For the same reason, the generalized price ratio for the 2-car alternatives and the constant terms in the 2-car utilities were omitted in the models for this life cycle. Because the number of licensed drivers in these segments is always 1 (if the number of licensed drivers is 0, only the 0-car and transit-to-work alternative is available and the observation is omitted from the data set), the cars per licensed driver and the inverse licensed drivers variable also were omitted.

The model for the market segment consisting of life cycle 5, those households with no primary worker, is also quite different from the typical joint model. Because no work trip is made by members in this life cycle, only car ownership is considered. All the variables describing the attributes of the alternative models for the work trips (cars per licensed driver, in-vehicle time, excess time/distance ratio, and CBD-work-place dummy) are omitted entirely. The set of alternatives is reduced from the 5 car-ownership and travel-mode-to-work alternatives to 3 options: 0, 1, and more than 1 car. Therefore, the 4 constants measuring the pure alternative effects are reduced to 2 constants; as before, the 0-car alternative is taken as a base. The omission of all level-of-service-to-work variables does not mean that transit level of service is excluded entirely from the market-segment-5 model. The generalized price ratio variables representing shopping accessibility still enter into the utility functions.

The remaining models are almost identical with 1 exception: The cars-per-licensed-driver variable was significant only for households with children. In the remaining life cycles, it is probable that the variation in the number of licensed drivers was insufficient to measure its effect on choice of travel mode to work. (The licensed-driver variable still entered into the model through its effect on the inverse licensed-driver variable, which only distinguishes among car-ownership levels but not travel modes to work.) Furthermore, experimentation with alternative measures leads to the formulation of a similar variable defined as the number of cars per licensed worker in the household.

Table 4 gives a summary of the final 7 models ultimately selected for the market segments. Each model is listed, the segment to which it applies is described, and the important characteristics of how the model specification differs from the general form described in the section on specification of the joint utility functions is given.

In addition to these models, a 1-in-5 sample of households in life cycles 1 through 4 was taken, and the model described in the section on specification of the joint utility functions was estimated. Because the behavioral process being modeled for this group is so different, members of life cycle 5 were not included in this pooled sample.

Coefficient estimates and "t-statistics" for the variables or constants of each of the models are given in Table 5. (The statistic in this computation is not distributed the same as the t-statistic is in ordinary least squares regression. However, for large data sets, both the t-statistic and this statistic approach a normal distribution.) Where a variable was not used, the entry is left blank. Table 6 gives the log likelihood function if all values were 0, $L^*(0)$, the log likelihood function for the actual estimates, $L^*(\beta)$, the number of observations, NOBS, and the total number of alternatives in excess of the number of observations for the entire data set, NCASES. (NCASES typically is given in reporting the results or choice model estimation because it reflects the number of degrees of freedom in the data set, that is, the number of alternatives in excess of 1 per observation.)

All coefficients in all of the models have the anticipated sign, and virtually all are significantly different statistically from 0 at the 90 percent confidence interval. Only 1 of the t-statistics is less than 1, and only 3 are less than 1.5. In each of these cases, the strong theoretical justification for including the variable overrode statistical considerations.

In no case is the question of whether the estimates, taken collectively, are significantly different from 0 in doubt. However, this is an extremely weak test because of the unreasonableness of the null hypothesis that all parameters are 0. Zero values imply that all the alternatives are equally likely. Nevertheless, failure to pass this test would indicate a serious problem with the models.

In all models in which the 2-car alternatives were available, the coefficients for the generalized price ratio variables were both negative, and the magnitude of the coefficient for the 2-car alternative was always greater than the corresponding value for 1 car. For a given household, both of these variables have the same value but each applies to a different utility. This indicates that any increase in car-shopping generalized price will produce a shift toward 0- and 1-car ownership. That is, the utility of 2-car ownership will decrease more than that of 1-car ownership. The 0-car alternative is unaffected because it was selected arbitrarily as the base against which the effect of the generalized price ratio is measured. Conversely, an increase in transit-shopping generalized cost will produce a shift toward multiple-car ownership.

In general, the coefficient of the natural logarithm of remaining income decreases sharply as one considers life cycles in the order in which a household might progress. Figure 3 shows this progression and a hand-fitted curve for the 8 models. This shift

Table 3. Coefficients appearing in utility functions.

Alternative	Constant	Cars/ Licensed Driver	Remaining Income	Single-Family-Dwelling Dummy	In-Vehicle Time	Excess Time/ Distance Ratio	Inverse Licensed Drivers	Generalized Price Ratio	CBD-Work-Place Dummy
0 car and transit taken to work	0	0	β_6	0	β_8	β_9	0	0	0
1 car and car taken to work	β_1	β_5	β_6	0	β_8	β_9	β_{10}	β_{11}	β_{13}
1 car and transit taken to work	β_2	0	β_6	0	β_8	β_9	β_{10}	β_{11}	0
2 cars and car taken to work	β_3	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{12}	β_{13}
2 cars and transit taken to work	β_4	0	β_6	β_7	β_8	β_9	β_{10}	β_{12}	0

Table 4. Summary of market-segment models.

Model	Segment	Description
1	1A	No 2-car alternatives considered; variables relating to these alternatives omitted
2	1B	Same as model 1
3	2A and 2B	Cars per licensed driver omitted
4	3A	Same as described in section on specification of joint utility functions; only half of data used
5	3B	Same as model 4
6	4A and 4B	Cars per licensed driver omitted
7	5	No mode to work considered; 4 alternative dummy variables reduced to 2; variables relating to mode attributes omitted

Table 5. Model coefficients and t-statistics.

	Model 1, Segment 1A		Model 2 Segment 1B		Model 3, Segments 2A and 2B		Model 4, Segment 3A		Model 5 Segment 3B		Model 6, Segments 4A and 4B		Model 7, Segment 5		Model 8, Segments Pooled	
Variable	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.	Coeff. Est.	t-Stat.
β_1	5.98	7.39	7.39	6.64	7.11	6.67	8.63	9.49	11.3	6.65	12.3	7.54			7.95	9.79
β_2	3.85	5.25	5.55	5.33	5.14	5.16	6.44	7.66	9.89	5.90	10.5	6.60			7.20	9.28
1 car													6.54	11.3		
β_3					9.13	7.07	10.1	8.22	13.9	7.22	15.9	8.32			9.58	9.89
β_4									11.0	5.62	15.4	6.59			7.35	7.59
2 cars													7.08	8.75		
β_5									1.92	3.80					1.92	5.29
No. cars/ licensed worker							0.249	1.19								
β_6	6.74	5.47	8.33	4.55	3.14	6.31	1.07	5.37	1.55	5.51	0.829	2.07	1.88	4.82	2.22	7.43
β_7[a]	0.578	1.71	0.232	0.440												
β_7[b]					0.778	5.11	1.50	9.55	1.32	9.86	0.970	6.27	0.734	2.27	1.26	9.49
β_8	-0.0117	-1.92	-0.131	-1.83	-0.00789	-1.54	-0.00658	-1.35	-0.017	-3.19	-0.0146	-2.72			-0.0120	-2.78
β_9	-0.0831	-2.05	-0.169	4.54	-0.0951	-2.12	-0.0879	-2.03	-0.100	-1.81	-0.107	-2.51			-0.0742	-2.19
β_{10}					-2.33	-3.39	-3.47	-5.58	-6.29	-6.33	5.17	-5.92	-2.04	-5.15	-3.59	-6.45
β_{11}	-5.65	-5.25	-6.59	-4.54	-3.38	-2.39	-5.14	-4.68	-5.11	-2.07	-7.07	-3.15	-5.31	-6.50	-6.46	-6.04
β_{12}					-4.72	-3.27	-6.49	-5.70	-6.45	-2.60	-9.35	-4.13	-6.11	-6.24	-8.17	-7.42
β_{13}	-0.821	-3.01	-0.982	-3.18	-1.33	-4.84	-1.40	-5.63	-2.15	-6.25	-1.58	-5.92			-1.40	-6.46

[a] 1 car. [b] 2 cars.

Table 6. Model log functions and other data.

Model	Segment	$L^*(0)$	$L^*(\beta)$	NOBS	NCASES
1	1A	-535.0	-325.5	487	974
2	1B	-608.6	-281.3	554	1,108
3	2A and 2B	-1,454	-893	1,092	3,317
4	3A	-2,053	-1,158	1,583	4,526
5	3B	-2,576	-1,295	1,982	5,846
6	4A and 4B	-1,973	-1,080	1,475	4,500
7	5	-718.2	-441.0	853	1,116
8	Pooled	-2,465	-1,392	1,899	5,514

in coefficient value reflects a change in how households perceive the marginal utility of added income for households with equal value of remaining income. Mathematically,

$$\text{Marginal utility of income} = \frac{\beta_{\text{remaining income}}}{\text{remaining income}}$$

where $\beta_{\text{remaining income}}$ is the coefficient of the variable. At first glance, this result might seem incorrect. One would think that households in their childbearing stage of life would have a higher marginal value of money than households in the other life cycle groups. To properly interpret this, however, one must recall that 2 households with equal values of the remaining income variable do not necessarily have equal income. The variable includes a deduction of $800/household member so that, given 2 households with equal remaining income, the larger household generally will have greater total income. Thus, if all else is equal, the households with children typically will have higher total incomes than those in other life cycles and therefore may not necessarily place a higher value on additional income. This argument still leaves open the question of why the marginal utility of money for a given value of remaining income should decrease so markedly. This effect can be attributed to an unmeasured socioeconomic attribute, wealth. Wealth is highly correlated with income and hence is highly correlated with remaining income. Furthermore, as households pass through life cycles 1 to 4, they typically accumulate wealth, in terms of both savings and property. Thus their car-ownership behavior shifts subtly; for example, they may no longer pay financing charges to purchase automobiles and therefore would perceive lower costs. For this reason, the households trade off the benefits and costs of car ownership in different ways depending on which life cycle they are in, and the relative weight of money decreases as the household accumulates wealth.

Another interesting effect is the value of the coefficient of the single-family-dwelling dummy variable. The way in which the coefficient varies across life cycles is shown in Figure 4. Note that this curve is unimodal; its peak is around life cycle 3. This is entirely consistent with the behavioral justification underlying this variable presented in the section on choice of variables. In that section, the effect of multi-family-dwelling ownership on car ownership was attributed to a number of factors one of which was the need to chauffeur children to a variety of activities. This behavioral factor, however, applies only to life cycle 3 because the households in the remaining life cycles have no children residing at home. Thus one would anticipate that the coefficient of the single-family-dwelling dummy would be higher for households in life cycle 3 than for households in the other life cycles.

FORECASTS FOR A TYPICAL HOUSEHOLD

To test how various types of transit service improvements might influence car ownership, we created a "typical suburban household" and applied the models to predict its expected automobile ownership in a variety of policy cases. (To measure aggregate impacts one needs to repeat the application of the model for a representative sample of households not just for a single typical household.) This household is not an average urban household; it is representative of a large group of suburban households residing in single-family dwellings outside the area now served by transit. It is this type of household toward which major transit extensions are most frequently directed. The typical household

1. Has a $13,500 income;
2. Resides in single-family dwelling;
3. Has a white-collar head of household who has a driver's license and works in downtown Washington, D.C.;
4. Has 2 children;

5. Has a spouse who has a driver's license; and
6. Lives in suburban location (Montgomery County) 5 miles (8.0 km) from workplace.

The choice probabilities for this typical household were evaluated for a base case and 3 other cases. Fare policy was held constant for all 4 cases. The level of service used was based on actual network travel times for a Washington, D.C., traffic zone in Montgomery County, Maryland, lying just outside the "10-mile square" (16-km square). The model used to predict the choice probabilities was model 5, which applies to all households with a white-collar breadwinner and with children under the age of 18 years residing at home. Table 7 gives some of the most significant findings resulting from this policy testing. Because the model predicts the probability with which various car-ownership levels will be selected, the expected or average car ownership is reported. Each of the successive levels of transit improvement influences car ownership. However, the greatest shift along the sequence occurs when transit of even relatively low quality is offered. Thus, the mere availability of transit service can have a substantial impact on car ownership. Furthermore, the model results indicate that, even with futuristic transit service offering extremely fast service for both work and shopping travel, approximately 65 percent of typical suburban families still will own 2 or more cars.

CONCLUSIONS

Car-ownership modeling typically has been paid little attention in the transportation planning process; however, the causal linkages between car ownership and the travel patterns in urban areas have long been recognized. The models described in this paper provide a behaviorally sound method of incorporating these linkages and provide a set of reliable, policy-sensitive forecasting tools by which car ownership and choice of travel mode to work can be predicted. In the course of study, 4 things have been demonstrated.

1. To deal with car-ownership and travel-mode-to-work decisions as jointly determined at the household level is feasible. Furthermore, by using disaggregate choice models, one can make effective use of readily available transportation planning data.
2. Car-ownership decisions are made on substantially different criteria by different households depending on their life cycles and occupations. The failure to adequately reflect these behavioral differences in a model will result in inaccurate and possibly misleading forecasts and will fail to adequately represent the distribution of changes in car ownership over various socioeconomic groups.
3. Transportation policy can have a small but measurable impact on the level of car ownership in an urban area. Furthermore, the effects of various aspects of the transportation system such as in-vehicle time, out-of-vehicle time and cost for the work trip of the household's primary worker as well as shopping-trip level of service can be isolated.
4. Introducing transit service to areas where it had been unavailable can have a marked effect on car ownership; the effect of improvements in existing transit is only marginal.

Future research efforts should be directed toward extending the scope of household decisions considered in a joint, disaggregate behavioral model. Research is currently under way to explore how residential-location and housing decisions can be incorporated into a joint model and thereby span the long-term transportation-related decisions a household makes. An improved understanding of these decisions and the role car ownership plays in determining the pattern of urban location should aid in the formulation of more behaviorally structured models for transportation planning.

Figure 3. Remaining income coefficients for various life cycles.

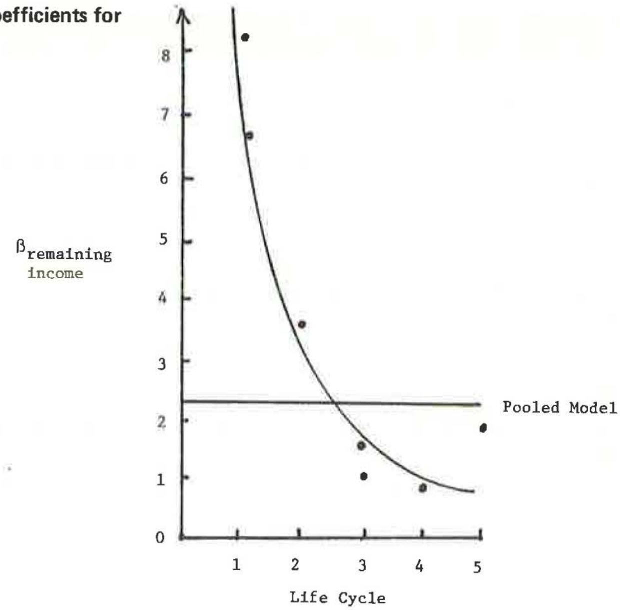


Figure 4. Type-of-dwelling dummy variable coefficients for various life cycles.

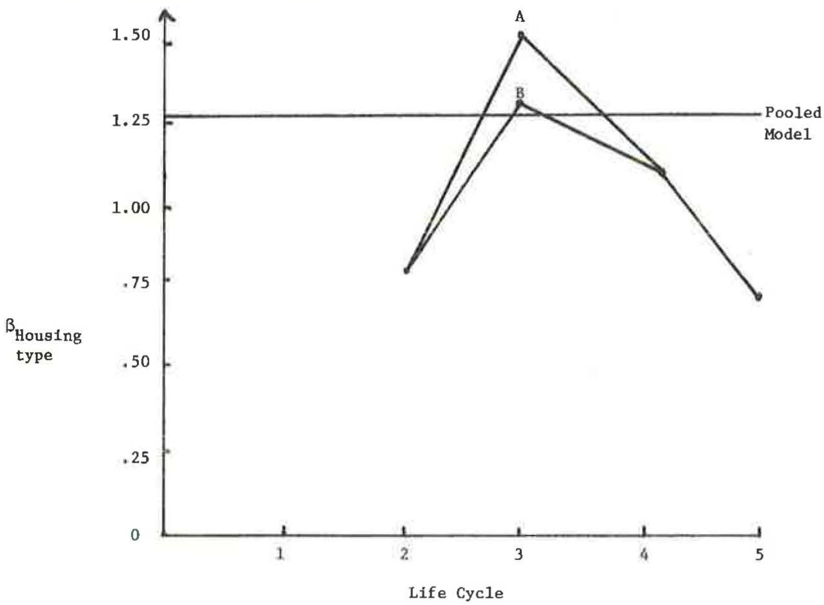


Table 7. Findings from policy testing of cases.

Case	Description	Expected Car Ownership	Change in Car Ownership From Base Case (percent)
Base	No transit available	1.825	—
1	Low-level transit	1.710	6.3
2	Good service (rail rapid transit) for work trips, case 1 transit service for shopping trips	1.660	9.0
3	Extremely high quality transit service (dense area, wide personal rapid transit)	1.622	11.1

ACKNOWLEDGMENTS

This paper is based on work performed at Cambridge Systematics, Inc., under contract to the Office of the Secretary and the Federal Highway Administration, U.S. Department of Transportation. However, the opinions expressed herein are solely ours and not necessarily those of the U.S. Department of Transportation.

We wish to acknowledge Richard L. Albright for his major programming and data-processing effort and William A. Jessiman, the principal-in-charge of this research project. We wish further to acknowledge the advice of Melvyn Cheslow and Edward Weiner, who both served as the U.S. Department of Transportation project monitors at different times during the study. Gregory Ingram, Daniel Brand, Robert Dunphy, Marvin L. Manheim, and Paul O. Roberts provided advisory support to the research. All errors and omissions, however, remain our sole responsibility.

REFERENCES

1. E. Aldana. Towards Microanalytic Models of Urban Transportation Demand. Department of Civil Engineering, Massachusetts Institute of Technology, PhD dissertation, 1971.
2. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, Massachusetts Institute of Technology, PhD dissertation, 1973.
3. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, Inc., 1972.
4. R. T. Dunphy. Transit Accessibility as a Determinant of Automobile Ownership. Highway Research Record 472, 1973, pp. 63-71.
5. M. J. Fertal, E. Weiner, A. J. Balek, and A. F. Sevin. Modal Split. Bureau of Public Roads, U.S. Department of Commerce, 1966.
6. P. Hoxie. Auto Ownership and the Long Run Transportation Decision. Department of Civil Engineering, Massachusetts Institute of Technology, MS thesis, 1970.
7. J. B. Lansing and G. Hendricks. Automobile Ownership and Residential Density. Institute for Social Research, Univ. of Michigan, 1967.
8. D. McFadden. The Revealed Preferences of a Government Bureaucracy. Institute of International Studies, Univ. of California, Berkeley, Technical Rept. 17, 1968.
9. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1973.
10. W. L. Mertz and L. B. Hamner. A Study of Factors Related to Urban Travel. Public Roads, Vol. 29, April 1957.
11. D. Mudarri. Mode Choice for Work Trips: A Key Urban Transportation Problem. Department of Economics, Univ. of Michigan, PhD dissertation, 1972.
12. W. Y. Oi and P. W. Shuldiner. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, Ill., 1962.
13. Peat, Marwick, Mitchell and Co. Implementation of the n-Dimensional Logit Model. Comprehensive Planning Organization, San Diego County, Calif., 1972.
14. P. L. Watson. Problems Associated With Time and Cost Data Used in Travel Choice Modeling and Valuation of Time. Highway Research Record 369, 1971, pp. 148-158.

DISCUSSION

Fred A. Reid, University of California, Berkeley

The paper presents a clear model structure for joint choice consistent with a behavorial theory of car ownership and travel mode to work. It also gives a convincing empirical test of the model with a comprehensive set of explanatory variables. The model is sufficiently simple that its joint-choice behavior and the influence of individual variables may be seen clearly, yet it still includes many of the variables influencing the choice set and a simple theory for the functional form of their entry. It seems to be a good base for building a more precise model of transportation ownership and modal choice. The paper also presents a valuable contribution to the evidence that stratification of a sample by life cycle or occupation or both can add to the explanatory power of a traveler model.

The brief theoretical arguments for the construction of the explanatory variables, though plausible and efficient because they include many important factors, are not supported sufficiently. The reader should expect some evidence or at least more comment about the forms considered or tested and rejected. The definition and form of entry of the remaining income variable left questions such as the constancy of car costs per unit and the logarithmic entry of all costs, particularly trip costs.

The stated assumption that ownership and choice of travel mode to work precede the non-work-trip decisions is partially inconsistent with the model structure and tests. Because travel mode to work neither enters the definition of the generalized non-work-trip price ratios nor is distinguished by a separate coefficient estimate for each mode to work, the opposite choice hierarchy is implied by the model. The non-work-trip choice, reflected through the price ratio variable, though conditioned on the car-ownership level, is established before the work-trip modal probability. Several other alternative choice hierarchies and corresponding model structures should be tested in any case because, for example, the originally intended precedence of ownership level over non-work-trip choice is not obviously true, especially for the second car.

The variable for number of licensed drivers is potentially dependent on behavioral choice, especially in good transit areas, and this possibly biases the corresponding coefficient.

The example of the effect of a policy change on a typical case by using the calibrated model is a helpful illustration of the application of and conclusions from this model.

In conclusion, the paper presents valuable initial results and a good base for further revelation of this type of joint traveler choice.

DISCUSSION

Joseph Berechman, Department of Civil Engineering,
State University of New York at Buffalo

The purpose of these remarks is to examine critically the 2 major hypotheses that underlie Lerman and Ben-Akiva's paper. This discussion will argue that these hypotheses are weak and unsubstantiated and will question the analysis that follows the hypotheses.

In brief, the hypotheses contend that the decision to purchase a car is a function of the individual's life cycle and is highly dependent on employment and residential decisions. More generally, Lerman and Ben-Akiva argue that choices of employment and residential location precede car ownership and are viewed as medium-term choices. This hierarchical structure of decision making strongly influences (albeit in a somewhat vague way) the selection and specification of variables for their model. However, when the major findings and conclusions are presented, they do not, in any meaningful way, reflect these assumptions. In particular, a change in the distribution of employment and residential locations is not shown to have significantly reduced (or increased) the expected level of car ownership. In general, I feel that these assumptions are questionable at best and most likely incorrect when applied to real-world data. For example, in 1970, there were approximately 54 million cars in all standard metropolitan statistical areas (SMSAs) (15); the average number of cars per household was 1.22 (and 1.42 cars/household outside the city center). The corresponding figures for 1960 were

34 million cars and 1.01 cars/household, which are changes of 56.7 and 20.5 percent respectively. As a first approximation, it can be argued that an average of more than 1 car/household indicates a tendency to own at least 1 car regardless of changes in the so-called long-term decisions. Furthermore, in the same period (1960-1970), 18 percent of the metropolitan labor force changed its employment location (54 percent outside city center) and about 45 percent of all households changed their residential location. These gross figures, although they conceal some fundamental trends, essentially reflect the suburbanization process in which households from cities and from outside the SMSAs (including other suburbs) migrated to suburbs. Given the state of public transportation systems, a precondition to such a movement is the availability of a private car purchased long before the actual relocation. For these reasons, the assumption regarding the household's sequential decision process is not supported by the facts.

It seems more plausible to argue that the decision to own an automobile (at least the first for households that own 2 or more) is mainly insensitive to locational decisions and rather should be ascribed to cultural, socioeconomic, and perhaps psychological factors. Indeed Lerman and Ben-Akiva's findings tend to support this contention. Table 7 indicates that the introduction of extremely high quality transit service to a hypothetical area where public transit services were previously unavailable will reduce expected car ownership from an average of 1.82 to 1.62. In other words, the desire to own a car is largely unaffected by the state of the transit system.

Another objection to Lerman and Ben-Akiva's hypotheses is based on the unavoidable feedback between short-term choices and long-term choices. For Figure 2, the block representing nonwork travel and employment and residential decisions should have been linked to reflect such interdependencies. In recent years, there has been a growing literature to indicate relationship between (a) residential amenities, environmental qualities, and neighborhood conditions (social and ethnic) and (b) employment and residential locational decisions. On a more general level, such relationships may be interpreted as implying a simultaneous decision process rather than a hierarchical one, and separation of choices into well-distinguished categories is, therefore, unjustifiable. [Interestingly enough, some authors (16) make the distinction between long- and short-term choices (residential location, place of work, and mode of travel to work) to characterize the behavior of low-income, unskilled laborers. Needless to say, the Lerner and Ben-Akiva study implicitly assumes middle-class white-collar households.]

The major assumptions that underlie Lerman and Ben-Akiva's paper are not supported by any available statistics nor by any known theory of urban spatial structure. Consequently, the methodology used in their study to explain car ownership is, in my opinion, rather questionable.

REFERENCES

15. Urban Transportation Factbook. American Institute of Planners, Part 1, 1974.
16. E. Banfield. The Unheavenly City. Little, Brown and Co., Boston, 1970.

AUTHORS' CLOSURE

Both Reid and Berechman raise a number of interesting points in their discussions of our paper. Some of these comments provide useful directions for future research; others raise some significant problems that were considered in the course of model development but were not resolved because of insufficient data or resources.

Reid notes the weakness of assuming a constant cost per automobile. In part, this simplification was required because the models focused only on the number of automobiles a household would own rather than on their age, make, power rating, or other aspects. Thus there is no way in the model to distinguish an old, dilapidated vehicle from a new, luxury-class car. Ultimately, when appropriate disaggregate data about households' choice of type of car become available, developing more detailed car-ownership models will be feasible.

Reid points out the partial inconsistency in formulating the generalized price measures of shopping trips; travel mode to work does not enter into the choice model used to generate these measures, and trip frequency is ignored. When the work described in our paper was performed, no known shopping-choice model included such interactions, and a reestimation of Ben-Akiva's model to allow for such effects was infeasible because of project resources. However, recent models described by Adler and Ben-Akiva (17) have explicitly included the effect of car-ownership and travel-mode-to-work decisions on non-work-trip frequency, destination, and modal choice. This should be a useful base for later extensions of our methodology.

Reid raises the question of the extent to which the number of licensed drivers is itself determined jointly with car ownership. In an attempt to explore the potential effect of this factor, some models that replaced the number of licensed drivers in the household with the number of household members old enough to drive were estimated. The results indicated that, at least within the groups tested, the coefficient estimates were not sensitive to the assumption that the number of licensed drivers is determined exogenously. This does not, however, exclude the possibility that, among specific household groups (most notably lower income and inner-city households), the effect that Reid hypothesized does not exist. Further work directed at understanding the car ownership of these particular groups should yield useful insights into such phenomena.

Berechman argues that the decision to own a first car is relatively insensitive to household-location decisions and is more closely linked to cultural, socioeconomic, and psychological factors. No attempt was made to measure either cultural or psychological factors, nor do we feel that the inclusion of variables measuring such factors is desirable in a model designed principally for forecasting the impact of alternative transportation policies. However, a wide range of important socioeconomic factors explicitly influence car ownership in our model, and the results given in Table 7 reflect their relative importance (compared with transit level of service) in determining first-car ownership.

Other criticisms raised in the discussions seem somewhat less well founded. Berechman suggests that the coefficients of the models are inconsistent because they represent car ownership and travel mode to work conditionally on location. He cites somewhat unconvincing aggregate statistics about the suburbanization process to argue the possibility that location decisions may be conditional on car ownership rather than the reverse.

As Ben-Akiva (18) points out, conditional-choice models produce consistent estimates of the utility-function parameters as long as the conditional structure is explicitly included in the utility specifications. Indeed, if this were not the case, the entire body of research into choice of travel mode to work would be invalidated. A conditional model, however, should be carefully applied because, in forecasting, some of the choices on which the forecasts are conditional may change. This is the reason that we term our model a medium-term model in the proposed-choice hierarchy. More comprehensive models that include jointly residential location, housing, car ownership, and travel mode to work have been developed recently by Lerman (19) and are more applicable to analyzing longer term policy impacts.

Berechman incorrectly concludes that nonwork travel and employment and residential location decisions are not linked in our model. In reality, exactly the opposite is true. This is precisely why the shopping generalized prices in our model are denoted by residence zone. Different places of residence have different shopping-trip patterns associated with them, and these trip patterns are linked directly to car-ownership choices.

The final significant area of concern raised in the discussions is the role of analytic theory in the development of behavioral models. We do not perceive our modeling approach and the more formal utility theory of Burns, Golob, and Nicolaidis as conflicting efforts. Explicit theories of how household members interact in making car-ownership and other decisions are important additions to understanding choice processes. We believe that such theories are the first steps in what will prove to be a fruitful research area.

However, strict theoretical approaches currently yield models that are not sufficiently rich in behavioral content to describe observed behavior in a way that is directly useful for analyzing transportation policies. This is the principal motivation for our using the household as the decision-making unit and including factors such as the number of licensed drivers, the number of licensed workers, and household size (as part of the remaining-income variable) in the model. These variables represent in an abstract way the outcome of the extremely complex bargaining process that takes place when a household makes a collective decision on car ownership. Including such variables is consistent with the basic goal of the study, the development of a useful analysis tool with which planners can assess the effect of transportation policy on car-ownership and travel-mode-to-work choice.

REFERENCES

17. T. J. Adler and M. E. Ben-Akiva. Joint-Choice Model for Frequency, Destination, and Travel Mode for Shopping Trips. Paper in this Record.
18. M. E. Ben-Akiva. Structure of Passenger Travel Demand Models. Transportation Research Record 526, 1974, pp. 26-42.
19. S. R. Lerman. Location, Housing, Automobile Ownership and Travel Mode to Work: A Joint Choice Model. Paper presented at the 55th Annual Meeting and to be published in a forthcoming Transportation Research Record.

THEORY OF URBAN-HOUSEHOLD AUTOMOBILE-OWNERSHIP DECISIONS

Lawrence D. Burns, Thomas F. Golob, and Gregory C. Nicolaidis,
Transportation and Urban Analysis Department,
General Motors Research Laboratories, Warren, Michigan

Automobile-ownership behavior is modeled as a function of socioeconomic factors and the availability and levels of service of public transportation systems. Decision makers are assumed to maximize their individual and household utilities within budget and time constraints. The benefits of increased mobility are weighed against the loss in other consumption attributable to ownership of 1 or more automobiles. Variables specific to residential and activity-center locations include the attractiveness of destinations served by public transit, the attractiveness of those not served by public transit, and respective travel times by automobile and transit. Estimation equations are developed through the introduction of functional forms for utility components and random utility terms representing variances in perception and taste and omitted factors. Multinomial logit models are used to define the probability of homogeneous groups of households choosing to own a specific number of automobiles. Calibrations are performed by using data from home-interview surveys and network simulations in the Detroit metropolitan area. Results are encouraging. All coefficients representing partial equilibrium market parameters are signed correctly and are significantly different from 0 where expected, and goodness-of-fit measures indicate acceptable model descriptive power.

•THAT an association exists among the number of automobiles available to an urban household, the total number of trips made by individual household members by various modes of transportation, the time of day of these trips, and their destinations (14, 28) is strongly supported by empirical evidence. That these empirical relationships are indicative of processes of the supply and demand of transportation services is axiomatic in well-known economic theories.

The research reported in this paper concerns the supply and demand relationships involved in household automobile-ownership decisions. These decisions are modeled in light of the accessibility of travel destinations by automobile and by alternative public transit modes, the service characteristics of these modes, and the income and demographic characteristics of the households themselves.

An earlier paper covering this research project proposed a theory for explaining automobile ownership in terms of generalized concepts of urban spatial-location factors (1). In this paper a model is developed from this theory for use as a tool in urban transportation planning procedures. First, theoretical considerations are addressed with the objective of accounting for the decisions of individual travelers within a framework of household-level automobile-ownership decisions. Second, the simplifying assumptions necessary before the theoretical model can be made operational with data from traditional transportation planning home-interview surveys and land use surveys are detailed. Third, estimation equations are formulated by postulating functional forms of model components and introducing stochastic terms.

Results are given from empirical calibrations of the model and initial tests of its hypotheses by using data from the transportation and land use study conducted in the Detroit metropolitan area (35). A multinomial logit model is employed to estimate the values of the coefficients for the postulated utility functions. The results are encouraging in that the estimated coefficient values of the model variables are in all cases correctly signed and have sufficiently small standard errors to reject the null hypotheses within traditional confidence limits. Traditional goodness-of-fit measures are at values that are quite acceptable for nonlinear estimation equations of the multinomial logit type. A new technique is introduced for assessing the goodness of fit of such nonlinear probabilistic choice models. The current model also looks good in terms of the various measures generated through use of this technique.

BACKGROUND

Forecasts of changes in automobile ownership resulting from changes in transportation systems or spatial activity patterns have been recognized as being important to the evaluation of costs and benefits for roadway or public transit investment levels. Most previous models developed for use in transportation planning have involved describing automobile ownership for spatially defined aggregations of households in the following terms: as a function of measures of residential density (10, 29, 30); as a function of family income (16, 29); and as a function of household socioeconomic and demographic characteristics (3, 22, 26, 31). The research efforts by Shindler and Ferreri (33) and by Dunphy (15) were the first known to us to explicitly incorporate transportation system characteristics into automobile-ownership forecasts. As discussed by Beckmann, Gustafson, and Golob (1), the current model differs significantly from these and other related efforts. The difference is primarily in terms of the use herein of an economic theory of decision-making behavior instead of the correlative relationships that characterize most other studies. Empirically, the definitions of accessibility and mobility also are different.

The current model, after it was further tested and refined, was judged to be an appropriate complement to traditional travel demand models. These traditional models usually are focused on the short-term demand and supply aspects of transportation such as choice of mode for a fixed destination trip, or, perhaps, choice of destination and time of trip given fixed residential location and automobile ownership.

Because this model is based on the use of individual households and individual travelers as the units of observation, the model is consistent with the class of short-term demand models referred to as disaggregate travel demand models (8, 34). Also, because the model is based on utility theory from microeconomics and welfare economics, it is intimately related to a series of models formulated by us and by others to explain various travel phenomena in terms of that theory (2, 9, 17, 18, 19, 21, 27).

BASIC THEORY

In Beckmann, Gustafson, and Golob (1), the decision-making behavior of households is postulated to be a result of individuals making trade-offs between the costs and benefits perceived to be associated with transportation-related alternatives. The idea that, if such household preferences are transitive and continuous, they can be represented by a numerical function called a utility function is well developed in economic theory (13).

The total utility to a household is defined here to be dependent on consumption of all goods, available leisure time, and travel to all destinations visited within a certain time period. The household increases its mobility with the purchase of an automobile but sacrifices other consumption; the decision to purchase is made when the utility to the household of the increased mobility exceeds the loss of utility of consuming other goods. A car is assumed to be a homogeneous good having a given fixed price. The extent to which an alternative to the car mode is available to satisfy the travel needs of the household is represented by the sets of destinations accessible by the alternative mode and by the travel times required to reach each destination. Assume that this alternative mode is the "best" public transit mode perceived to be available to a decision maker. Specifically, the utility to a household not owning an automobile is a function of income (representing all consumption), nontravel time, and trips to the set of destinations accessible by the modes of transportation that are alternatives to travel by personal automobile:

$$U_i^0 = U\left(y,\ T - \sum_{k \in D_0} r_{ik}X_{ik},\ X_{ik}\right) \tag{1}$$

where

U_i^0 = utility to a household without a car at location i,
y = total disposable income of the household,
T = available household leisure time,
r_{ik} = travel time from the household at location i to destination k by alternative mode,
X_{ik} = number of trips the household makes to destination k by using alternative mode, and
D_0 = set of all destinations accessible to the household by alternative mode.

Assume that when an automobile is available household members will make an insignificant number of trips by the alternative mode of transportation. Total utility for automobile-owning households then is specified as a function of the number of trips to destinations accessible by automobile:

$$U_i^1 = U\left(y - p,\ T - \sum_{k \in D_1} s_{ik}Z_{ik},\ Z_{ik}\right) \tag{2}$$

where

U_i^1 = utility to a household with a car,
p = annual cost of owning and operating an automobile (assumed to be independent of the number of trips),
s_{ik} = travel time for the household at location i to destination k by automobile,
Z_{ik} = number of trips for the household to destination k by automobile, and
D_1 = set of destinations accessible to the household by automobile.

The ownership of a single automobile is advantageous to the household whenever

$$\max_{Z_k} U_i^1 > \max_{X_k} U_i^0 \tag{3}$$

The household thus assesses the maximum utilities that can be derived from making the most out of travel by automobile or travel by alternative mode. It compares these utilities and then makes its decisions regarding automobile purchase. Changes in income, automobile or alternative mode costs, automobile or alternative mode travel times, or accessibilities call for reassessments. Note that travel times, costs, and accessibilities to all destinations, visited or not visited, are taken into account here because these factors cause readjustments in utility-maximizing travel patterns. Also note that residential location is a very important factor in this model because change in residential location dictates changes in each of the explanatory variables.

This theory makes explicit a proposed relationship between automobile-ownership decisions and transportation system characteristics. Specification of functional forms for the utilities and the introduction of proxy variables are necessary before the theory can be implemented on currently available household home-interview and transportation system data. Also, to assess total automobile ownership, one must extend the theory to include decisions about additional automobiles. And, in both the single-car and multiple-car cases, one must consider the interactions of individual household members with respect to their travel needs and desires and their roles in automobile-ownership decisions.

MULTIPLE TRAVELERS AND MULTIPLE AUTOMOBILES

In looking at the increased household mobility due to the purchase of an automobile, the previously formulated theory considers the change in utility to the entire household as the result of the purchase. To properly assess the utility the household obtains from an automobile, one must consider the travel utilities of each individual trip maker in the household. Thus one must develop postulates about the way individual trip makers in the household interact in their usages of 1 or more family automobiles. Briefly, there are 3 postulates.

1. Each trip maker in the household maximizes his or her individual travel utility independently of the other trip makers in the household.
2. When a household purchases an automobile, 1 trip maker in the household is considered to have the exclusive use of this automobile. Thus the utility of 1 automobile to a household is reflected in the utility of the automobile to 1 trip maker in the household.
3. Trip makers who do not have the use of an automobile are indifferent about whether to use public transit (with greater travel times and limited accessibility) or postpone trips until the automobile becomes available. (For purposes of simplifying terminology, refer to travel by automobile when it becomes available as travel by alternative mode.)

Based on the 3 simplifying assumptions, the total utility of a household can be viewed as the aggregation of the travel utilities of individual trip makers within the household and the household-level residual-consumption term. The travel utility of each trip maker within the household is a function of his or her available leisure time and travel to all destinations visited within a specified time period.

In general, the set of destinations accessible by automobile differs from the set of destinations accessible by alternative mode of transportation. Also automobile travel times usually differ from the travel times of the alternative mode; walking and waiting times for public transit service are reflected in additional, perhaps more heavily weighted, time penalties for the alternative mode. Travel utilities of individual trip makers thus are dependent on whether the trip maker has the use of an automobile or whether he or she must rely on the alternative mode of transportation.

Specifically, the net travel utility to an individual j from travel by the alternative mode of transportation is considered to be separable into 2 components: net leisure time and satisfaction of travel purposes (the household location subscripts i are dropped here for simplification):

$$TU_j^0 = U\left(T_j - \sum_{k \in D_0} r_k X_{kj},\ X_{kj}\right) \tag{4}$$

where

TU_j^0 = net utility to an individual j from travel by alternative mode of transportation,
T_j = total leisure time of individual j, and
X_{kj} = number of trips individual j makes to destination k when relying on alternative mode.

Similarly, the net utility to an individual j from travel by an automobile is:

$$TU_j^1 = U\left(T_j - \sum_{k \in D_1} s_k Z_{kj},\ Z_{kj}\right) \tag{5}$$

where

TU_j^1 = net utility to an individual j from travel by exclusive use of an automobile, and
Z_{kj} = number of trips individual j makes to destination k when using an automobile.

If next the assumption is made that the total utility to a household at location i is additive with respect to the utility from income available for other consumption and total household travel utility, then

$$U_{im}^n = \phi\left(y - \sum_{\ell=1}^{n} p_\ell\right) + \sum_{j=1}^{n} TU_{ij}^1 + \sum_{j=n+1}^{m} TU_{ij}^0 \tag{6}$$

where

ϕ = functional form of the residual consumption contribution to utility, and
U_{im}^n = total utility of an m-trip-maker household at location i owning n automobiles.

Thus by substituting equations 4 and 5 in equation 6,

$$U_{im}^n = \phi\left(y - \sum_{\ell=1}^{n} p_\ell\right) + \sum_{j=1}^{n} U\left(T_j - \sum_{k \in D_0} r_{ik} X_{ikj},\ X_{ikj}\right)$$

$$+ \sum_{j=n+1}^{m} U\left(T_j - \sum_{k \in D_1} S_{ik} Z_{ikj},\ Z_{ikj}\right) \tag{7}$$

This is the foundation for the model specifying the conditions under which automobile-ownership decisions are undertaken. To test the model hypotheses however, one must specify mathematical forms for the ϕ and U functions and develop estimation equations.

FUNCTIONAL FORMS

A utility function that is logarithmic in terms of both time and trips was selected on the basis of its theoretical properties (such as its property of "diminishing marginal utility") and on the basis of its success in describing spatial interactions (19). This functional form also is related intimately to entropy formulations of trip distributions (41) and other transportation phenomena (42). Equations 4 and 5 then can be written respectively as follows (for simplicity the household location subscript has been dropped):

$$TU_j^0 = a_t \log\left(1 + T_j - \sum_{k \in D_0} r_k X_{kj}\right) + \sum_{k \in D_0} a_k \log(1 + X_{kj}) \tag{8}$$

$$TU_j^1 = a_t \log\left(1 + T_j - \sum_{k\epsilon D_1} s_k Z_{kj}\right) + \sum_{k\epsilon D_1} a_k \log(1 + Z_{kj}) \tag{9}$$

where

a_k = the value or attraction of destination k, and
a_t = the value of leisure time.

Each trip maker in the household maximizes his or her individual utility by adjusting number and distribution of trips. This maximization is performed independently of other members of the household in all matters not related to the availability of the household automobile or automobiles. In this process, the maximum household utility is not necessarily achieved.

Finding the maximizing solution for TU_j^0 is facilitated through the introduction of a new set of variables:

$$\left.\begin{aligned} \zeta_{kj} &= 1 + X_{kj} \\ R_j^0 &= 1 + T_j + \sum_{k\epsilon D_0} r_k \\ A_0 &= a_t + \sum_{k\epsilon D_0} a_k \end{aligned}\right\} \tag{10}$$

Maximizing equation 8 with respect to X_{kj} then implies

$$\frac{a_k}{1 + X_{kj}} = \frac{a_t r_k}{1 + T_j - \sum_{k\epsilon D_0} r_k X_{kj}} \tag{11}$$

This equation may be solved for ζ_{kj}:

$$\zeta_{kj} = \frac{a_k R_j^0}{A_0 r_k} \tag{12}$$

Similarly, let

$$\left.\begin{aligned} \eta_{kj} &= 1 + Z_{kj} \\ R_j^1 &= 1 + T_j + \sum_{k\epsilon D_1} s_k \\ A_1 &= a_t + \sum_{k\epsilon D_1} a_k \end{aligned}\right\} \tag{13}$$

Maximizing equation 5 with respect to Z_{kj} and solving for η_{kj} yield

$$\eta_{kj} = \frac{a_k R_j^1}{A_1 s_k} \tag{14}$$

Substituting ζ_{kj} and η_{kj} into the utility equations 8 and 9 and simplifying yield

$$\underset{X_{kj}}{\text{Max}} \, (TU_j^0) = a_t \log \frac{a_t R_j^0}{A_0} + \sum_{k \in D_0} a_k \log \frac{a_k R_j^0}{A_0 r_k} \tag{15}$$

$$\underset{Z_{kj}}{\text{Max}} \, (TU_j^1) = a_t \log \frac{a_t R_j^1}{A_1} + \sum_{k \in D_1} a_k \log \frac{a_k R_j^1}{A_1 s_k} \tag{16}$$

Distinguishing among the individual trip makers within a household is not within the scope of this theory. Thus

$$\text{Max}\, (TU_1^0) = \text{Max}\, (TU_2^0) = \ldots = \text{Max}\, (TU^0) \tag{17}$$

$$\text{Max}\, (TU_1^1) = \text{Max}\, (TU_2^1) = \ldots = \text{Max}\, (TU^1) \tag{18}$$

The travel-utility-maximizing form of the utility function given in equation 7 for a household at location i then becomes

$$U_{im}^n = \phi\left(y - \sum_{\ell=1}^{n} p_\ell\right) + n \, \text{Max}\, (TU_i^1) + (m - n) \, \text{Max}\, (TU_i^0) \tag{19}$$

where Max (TU^1) and Max (TU^0) are given by equations 15 and 16 and p_ℓ = annual cost to a household of owning and operating the ℓth automobile.

Hypothesizing that the utility of all other consumption has a similar logarithmic "diminishing marginal utility" functional form, one can write equation 19 without any general terms as

$$U_{im}^n = b_c \log\left(y - \sum_{\ell=1}^{n} p_\ell\right) + b_t \left[n \, \text{Max}\, (TU_i^1) + (m - n) \, \text{Max}\, (TU_i^0)\right] \tag{20}$$

where

b_c = additive parameter adjusting the scale of the utility of residual consumption to that of overall utility, and

b_t = additive parameter adjusting the scale of the utility of travel to that of overall utility.

CONDITIONAL AUTOMOBILE-OWNERSHIP DECISION

Assume that the maximum number of automobiles the household will own does not exceed the number of driver-aged trip makers in that household m. Now whenever

$$U_{im}^{\eta} > U_{im}^{\theta} \quad \text{for } \theta \neq \eta,\ \theta = 0, 1, \ldots, m \tag{21}$$

the household will purchase η automobiles. Specifying this condition in terms of the indirect utility function (equation 20) means

$$0 < U_{im}^{\eta} - U_{im}^{\theta} = b_o\left[\log\left(y - \sum_{\ell=1}^{\eta} p_\ell\right) - \log\left(y - \sum_{\ell=1}^{\theta} p_\ell\right)\right]$$

$$+ b_t\,[\eta \text{ Max }(TU_i^1) + (m - \eta) \text{ Max }(TU_i^0) - \theta \text{ Max }(TU_i^1)$$

$$- (m - \theta) \text{ Max }(TU_i^0)] \tag{22}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$. Simplifying means

$$0 < U_{im}^{\eta} - U_{im}^{\theta} = b_o\left[\log\left(y - \sum_{\ell=1}^{\eta} p_\ell\right) - \log\left(y - \sum_{\ell=1}^{\theta} p_\ell\right)\right]$$

$$+ b_t(\eta - \theta)\,[\text{Max }(TU_i^1) - \text{Max }(TU_i^0)] \tag{23}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$.

Substituting the indirect maximum travel utilities of equations 15 and 16 into this condition for ownership of η automobiles yields

$$0 < U_{im}^{\eta} - U_{im}^{\theta} = b_o\left[\log\left(y - \sum_{\ell=1}^{\eta} p_\ell\right) - \log\left(y - \sum_{\ell=1}^{\theta} p_\ell\right)\right]$$

$$+ b_t(\eta - \theta)\left(a_t \log \frac{a_t R^1}{A_1} - a_t \log \frac{a_t R^0}{A_0}\right.$$

$$\left. + \sum_{k \in D_1} a_k \log \frac{a_k R^1}{A_1 s_k} - \sum_{k \in D_0} a_k \log \frac{a_k R^0}{A_0 r_k}\right) \tag{24}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$.

However, by using the definitions of A_1, A_0, R^1, and R^0 specified earlier, one can simplify equation 24:

$$0 < U_{im}^{\eta} - U_{im}^{\theta} = b_c\left[\log\left(y - \sum_{\ell=1}^{\eta} p_\ell\right) - \log\left(y - \sum_{\ell=1}^{\theta} p_\ell\right)\right]$$

$$+ b_t(\eta - \theta)\left(A_0 \log\frac{A_0}{R^0} - A_1 \log\frac{A_1}{R^1}\right.$$

$$\left. + \sum_{k\epsilon D_0} a_k \log\frac{r_k}{s_k} + \sum_{k\epsilon D_1 - D_0} a_k \log\frac{a_k}{s_k}\right) \tag{25}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$. The set $D_1 - D_0$ is the set of destinations accessible by automobile but not accessible by transit; it is a simplified notation for the set represented by the intersection of D_1 with the complement of D_0.

Strictly speaking, A_1, A_0, R^1, and R^0 depend on household location i. However, these terms are rather insensitive to exact location; they will be treated as constants in the initial tests of the model hypotheses given herein. Thus,

$$0 < U_{im}^{\eta} - U_{im}^{\theta} = \alpha + b_c\left[\log\left(y - \sum_{\ell=1}^{\eta} p_\ell\right) - \log\left(y - \sum_{\ell=1}^{\theta} p_\ell\right)\right]$$

$$+ b_t(\eta - \theta)\left(\sum_{k\epsilon D_0} a_k \log\frac{r_k}{s_k}\right)$$

$$+ b_t(\eta - \theta)\left(\sum_{k\epsilon D_1 - D_0} a_k \log\frac{a_k}{s_k}\right) \tag{26}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$ where $\alpha = (\eta - \theta)\left(A_0 \log\frac{A_0}{R^0} - A_1 \log\frac{A_1}{R^1}\right)$.

The second term on the right side of equation 26 indicates the difference in utility to the household from consumption of all other goods when the household owns θ automobiles compared to η automobiles. The third term on the right indicates the utility $(\eta - \theta)$ that trip makers in the household receive from travel time savings when traveling by automobile to destinations that can be reached by transit (destinations in set D_0), and the fourth term on the right indicates the utility $(\eta - \theta)$ that trip makers receive from potential travel to destinations not accessible by public transit but accessible by automobile. The third and fourth terms together comprise the travel component.

ESTIMATION EQUATIONS

The model of automobile ownership developed in this paper is, of course, a simplification of reality. It does not take into account all of the differences in the tastes and perceptions of individual decision makers. Also data are not available to measure all of the stimuli that decision makers might consider important. Thus the values of the utility differences given in equation 26 are random variables across samples of households.

By letting ϵ_m^{η} represent an unobserved random variable containing excluded decision factors and individual differences, one can specify the probabilistic utility function for an m-trip-maker household owning η automobiles as

$$RU_m^\eta = f(U_m^\eta, \epsilon_m^\eta) \tag{27}$$

This function generally can be scaled so that it can be written as the sum of the deterministic component and the random component ϵ_m^η:

$$RU_m^\eta = U_m^\eta + \epsilon_m^\eta \tag{28}$$

The probability that a sampled household finds ownership of η automobiles more advantageous than ownership of any other number of automobiles θ can then be denoted by

$$P_m^\eta = \text{Prob}\,[(U_m^\eta + \epsilon_m^\eta) > (U_m^\theta + \epsilon_m^\theta)] \tag{29}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$ where P_m^η = probability that the household will desire to own η automobiles. Equation 29 can then be rewritten as

$$P_m^\eta = \text{Prob}\,[(\epsilon_m^\theta - \epsilon_m^\eta) < (U_m^\eta - U_m^\theta)] \tag{30}$$

for $\theta \neq \eta$, $\theta = 0, 1, \ldots, m$.

The problem of developing estimation equations for the current automobile-ownership model is now one of specifying a distribution for the probabilistic components in equation 30. The one chosen here is the reciprocal exponential distribution; random variables are assumed to be distributed independently as

$$\text{Prob}\,(\epsilon_m^\theta \leq w) = \exp(-e^{-w}) \tag{31}$$

Probability equation 29 then takes the form

$$P_m^\eta = \frac{\exp(U_m^\eta)}{\sum_{\theta=0}^{m} \exp(U_m^\theta)} = \frac{1}{\sum_{\theta=0}^{m} \exp(U_m^\theta - U_m^\eta)} \tag{32}$$

This function is termed the conditional or multinomial logit function. It was first developed systematically by Gurland, Lee, and Dolan (20). A more general formulation in terms of choice behavior is provided by McFadden (24, 25). Other general developments are provided by Bloch and Watson (5), Bock (6), and Theil (37, 38). Applications of the functional form to transportation mode, destination, and trip-time choice were made by Rassam, Ellis, and Bennett (32) and Charles River Associates (9).

The probability that an m-trip-maker household decides to own η automobiles thus is specified by equation 32; the $U_m^\eta - U_m^\theta$ utility differences for the alternative choices of $\theta = 0, 1, \ldots, m$, $\theta \neq \eta$ are given by equation 26. By observing the actual choices made by households when they are faced with various values of the independent variables, one can make estimates of the constant α and the b_o and b_t coefficients in equation 26 through the application of a maximum-likelihood procedure proposed by McFadden (24) and implemented by Manski (23). Statistics generated in this first empirical test of the model hypotheses are the subject of the remainder of this paper.

DATA

The data employed in the empirical test were obtained from individual household observations and network simulation results from the 1965 Detroit Transportation and Land Use Study (TALUS) (35). The variables to be used in the following analysis are defined as follows:

y = disposable income of the household (calculated by subtracting estimated taxes from respondents' reported gross income);
$p_{i\ell}$ = annual cost to a household at location i of owning and operating the ℓ th automobile (assumed to be independent of the number of trips);
r_{ik} = travel time from household at location i to destination k by public transit [generated from the 1965 Detroit transit network by using the Urban Transportation Planning System (UTPS) (40) including walking, waiting, transfer, and running times];
s_{ik} = travel time from household at location i to destination k by automobile (taken directly from network simulation results generated in the TALUS study);
D_i^0 = set of destinations accessible to an individual trip maker at location i by public transit (a destination k was considered accessible by transit if it could be reached within a specified period of time by transit);
D_i^1 = set of destinations accessible to an individual trip maker at location i when this trip maker has the exclusive use of an automobile (a destination k was considered accessible by automobile if it could be reached within a specified period of time by automobile);
f = number of automobiles a household owned in 1965; and
a_{ik} = attraction of destination k to a household at location i.

Individual household observations were taken from the 4 percent survey of all households within the Detroit urbanized area, and the origin and destination subscripts (i and k) correspond to traffic analysis zones. The limit of the study spatial area is defined by the inclusion of all traffic analysis zones in TALUS that are located within the Detroit urbanized area as defined by the 1960 Census (34).

For the purposes of this initial test of the theory, households with the same number of driver-aged trip makers m were assumed to be homogeneous with respect to their automobile-buying behavior. The data were sorted into 3 sets:

1. Households with 1 driver-aged trip maker,
2. Households with 2 driver-aged trip makers, and
3. Households with 3 or more driver-aged trip makers.

Certain assumptions were made in preparing the data. These assumptions are the subjects of model-sensitivity analyses.

1. Automobile ownership cost $p_{i\ell}$ is constant for all automobiles and is independent ($= p_\ell$) of location i; an average figure of \$1,000 (1965 prices) was selected on the basis of automobile cost data provided by Botzow (7).
2. A destination is a member of the set D_i^0 if it can be reached from i in 60 min by transit; a destination is a member of the set D_i^1 if it can be reached from i in 60 min by automobile; $D_i^0 \subset D_i^1$.
3. The attraction of a destination k is independent of the location of the household ($= a_k$) and is given by

$$a_k = \frac{E_k + P_k}{\sum_k (E_k + P_k)} \tag{33}$$

where

E_k = total employment at destination k, and
P_k = total population residing at destination k.

The summation in the denominator is over all traffic analysis zones in the Detroit urbanized area.

EMPIRICAL RESULTS

The multinomial logit estimation equation 32 for the utility differences of equation 26 was calibrated for each of the 3 population segments (1 trip maker, 2 trip makers, and 3 or more trip makers). Random samples of between 500 and 1,000 households in each segment were selected from the set of all households responding to TALUS (35). These samples were structured to obtain an approximately equal number of observed households that chose each of the alternatives; this structuring resulted in a higher sampling rate for 0-car households. Samples that were held out were employed for assessing goodness of fit of the models.

Two different forms of the model were calibrated for each of the 3 segments. In the simplest (choice-abstract) form, the utility scale weights b_c and b_t are assumed to be independent of the choice alternatives. This is the usual assumption underlying use of multinomial logit functions in modeling travel-mode choice. It is the only form of the model applicable to the binomial 1-trip-maker case in which only the 2 choices of 0 car and 1 car are theorized to be relevant.

The second and more complicated form of the model for the 2-trip-maker and 3-or-more-trip-maker segments is developed by assuming that the b_c and b_t coefficients are specific to each sequential choice alternative. That is, it is proposed that the weights people place on consumption and travel components of utility may be different when they are evaluating the 0-car and 1-car choice alternatives than when they are evaluating the 1-car and 2-car alternatives and so forth. In the calibration of this form of the model, freedom thus is incorporated to estimate as many b_c and b_t coefficients as there are choice alternatives minus 1. Referring to these coefficients as b_t^i and b_c^i, where i denotes the choice between i and 0 automobiles, is convenient.

For purposes of clarifying the previous arguments, the following estimation equations can be written for the most general case of 3 or more trip makers in which 4 choices (0 car, 1 car, 2 cars, and 3 cars) are relevant (P_m, probability that a household made up of m driver-aged trip makers will decide to own η automobiles, takes the form of equation 32 for $\eta = 0$):

$$\left.\begin{aligned}
U_m^1 - U_m^0 &= a_0 + b_c^1[\log(y - p) - \log y] + b_t^1\left(\sum_{k\epsilon D_0} a_k \log\frac{r_k}{s_k} + \sum_{k\epsilon D_1 - D_0} a_k \log\frac{a_k}{s_k}\right)\\
U_m^2 - U_m^0 &= a_0 + b_c^2[\log(y - 2p) - \log y] + b_t^2\left[2\left(\sum_{k\epsilon D_0} a_k \log\frac{r_k}{s_k}\right.\right.\\
&\qquad \left.\left. + \sum_{k\epsilon D_1 - D_0} a_k \log\frac{a_k}{s_k}\right)\right]\\
U_m^3 - U_m^0 &= a_0 + b_c^3[\log(y - 3p) - \log y] + b_t^3\left[3\left(\sum_{k\epsilon D_0} a_k \log\frac{r_k}{s_k}\right.\right.\\
&\qquad \left.\left. + \sum_{k\epsilon D_1 - D_0} a_k \log\frac{a_k}{s_k}\right)\right]
\end{aligned}\right\} \quad (34)$$

Estimation equations for the other 2 population segments merely eliminate 1 or 2 of the utility differences in equation 34. The choice-abstract form of the model will have $b_c^1 = b_c^2 = b_c^3 = b_c$ and $b_t^1 = b_t^2 = b_t^3 = b_t$. The constant a_0 includes the constant $\propto$ of equation 26 and the mean of the unobserved additive random term.

The results of the logit estimations for the 3 population segments for the choice-abstract form of the model are given in Table 1. The ratios are asymptotically distributed as t-statistics in a linear model by Theil (38) and thus can be used with large samples (as in the current case) to evaluate the probability that the coefficient estimates are in actuality 0. In the -2 log λ statistics for each model, λ is the ratio of the initial likelihood for the model with all coefficients as 0 and the final (maximum) likelihood with the coefficients as listed. This statistic, the so-called likelihood ratio statistic, has an approximate χ^2 distribution. Therefore, it can be used to test the joint hypothesis that the data are a result of processes inconsistent with the proposed theory (that is, the joint hypothesis: $b_c = b_t = 0$).

The -2 log λ statistics in Table 1 indicate firm rejections of the joint hypothesis $b_c = b_t = 0$ for all population segments. Moreover, the estimated coefficient values of the variables of the models derived from the theory are in all cases correctly signed, and the t-statistics indicate firm rejections of the hypothesis $b_c = 0$ or $b_t = 0$ for all population segments. These are encouraging initial results for internal tests of the current automobile-ownership theory, but further goodness-of-fit assessments are called for.

The results of the logit calibrations for the choice-dependent form of the model were judged to be inferior to the results of the choice-abstract form. The criteria for this judgment were the χ^2 statistics for the overall models and the t-statistics for the individual model coefficients. The χ^2 statistics did not significantly increase in spite of the increase in the degrees of freedom in the choice-dependent form. The hypothesis that individual b_c^i and b_t^i values were insignificantly different from 0 was accepted for some coefficients at as high as the 10 percent confidence limit.

Sensitivity analyses were performed by using 3 different values of the annual automobile cost parameter p; the values were \$750, \$1,000, and \$1,500. The model based on the original \$1,000 estimate was judged to be best on the basis of having the greatest t-statistic absolute values for the coefficients. The model based on the \$750 estimate was only slightly different from the chosen model, and the model based on the \$1,500 price was definitely inferior. Sensitivity analyses on other direct and definitional model parameters are within the realm of further research.

EVALUATION OF EMPIRICAL RESULTS

The statistics discussed in the preceding section provide insufficient information for assessing the goodness of fit of probabilistic or quantal choice models such as multinomial logit. The single overall measure of fit, the likelihood ratio test, is a rather insensitive test because of the questionable validity of the null hypothesis $b_c = b_t = 0$. Also model significance generally becomes easier to obtain with this test as sample size increases.

Table 1. Results of logit estimations.

	α_0		b_c		b_t		
Population Segment (number of trip makers)	Constant	Ratio of Estimated Coefficient Value to Estimated Standard Error	Utility Weight	Ratio of Estimated Coefficient Value to Estimated Standard Error	Utility Weight	Ratio of Estimated Coefficient Value to Estimated Standard Error	-2 log λ[a]
1	-3.17	-6.84	1.34	13.2	1.28	9.05	323
2	-0.71	-3.78	8.77	12.0	0.48	12.3	354
3 or more	-0.33	-1.80	6.81	9.70	0.34	10.2	144

[a]3 degrees of freedom.

Other measures of goodness of fit have been developed that are comparable to the coefficient of determination r^2 in linear models. Use of the r^2 formula directly to measure proportion of variance explained is inappropriate because the dependent observations by definition lie on the asymptotes of the logit function. It can be shown that the maximum value of r^2 for continuous independent variables cannot be equal to 1.0 as it is in the continuous dependent variable case, and this maximum value cannot be determined deductively. More important, r^2 is a measure of linearity, and this is not a correct criterion for the logit model.

Two statistics have been developed that are attempted analogies to r^2. Both are referred to as "pseudo r^2" and usually are denoted by ρ^2. The first, ρ_1^2, applied by Cragg (11), makes assumptions about the error term distribution resulting in the following mathematical expression:

$$\rho_1^2 = 1 - \exp\{-2[L^*(\hat{\theta}) - L^*(0)]/N\} \tag{35}$$

where

$L^*(\hat{\theta})$ = value of log likelihood function for vector of estimated coefficients $(\hat{\theta})$,
$L^*(0)$ = value of log likelihood function with all coefficients equal to 0, and
N = number of observations.

A second pseudo r^2 statistic, ρ_2^2, applied by McFadden (24) and discussed by Ben-Akiva (4), is equal to the ratio of explained log likelihood over total log likelihood and is expressed as follows [$L^*(\hat{\theta})$ and $L^*(0)$ are defined as they are in equation 35]:

$$\rho_2^2 = 1 - \frac{L^*(\hat{\theta})}{L^*(0)} \tag{36}$$

Values of these pseudo r^2 statistics for the automobile-ownership models of Table 1 are tabulated as follows:

Population Segment (number of trip makers)	ρ_1^2	ρ_2^2
1	0.259	0.216
2	0.428	0.255
3	0.348	0.155

Four problems are associated with using either pseudo r^2 statistic (ρ_1^2 or ρ_2^2) as a measure of goodness of fit.

1. No well-developed distribution theory or associated statistic exists (such as an F-statistic) for either measure to permit an assessment of the statistical significance of the measure.
2. Neither statistic is comparable between models with different functional forms.
3. Maximum value for each statistic is not defined clearly.
4. Neither statistic provides an intuitive interpretation (comparable to percentage of variance explained) of goodness of fit.

In an effort to improve the understanding of the goodness of fit of probabilistic choice models, we propose a new technique. To describe this technique in detail and discuss its purported strengths and weaknesses are not within the scope of this paper. A brief explanation is advanced, instead, and 1 example application is presented with the objective of providing insight into the performance of the models derived from the current automobile-ownership theory. Research publications directed to the specific subject of the goodness-of-fit evaluation technique are anticipated.

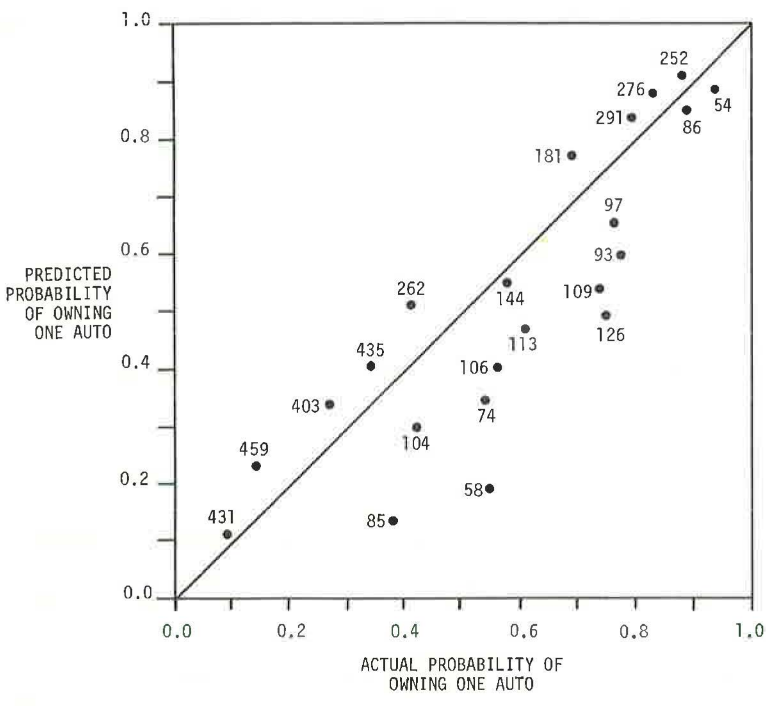
The proposed technique entails the formation of homogeneous groups of individual observations belonging to cells (classes) defined jointly by ranges of the independent (explanatory) variables of the model. Individuals in every cell, to an extent, we assumed to be faced with identical stimuli when evaluating their choice alternatives. Thus a hypothetical observation described by the mean values of the independent variables for all observations in a particular cell can be considered a representative observation for all the observations in this cell. The predicted probabilities for this particular representative observation then can be computed by using the estimated choice models. These predicted probabilities for each choice alternative finally can be compared to the probabilities given by the proportion of individuals in a particular cell that choose that alternative.

Figure 1 shows a typical scatter plot of predicted versus actual proportions of households owning no automobiles for a hold-out sample from the 1-trip-maker population segment. Each point represents a unique cell; the number corresponding to each point reflects the number of households in that cell. Ideally, all of these points should lie along the 45-deg line.

To compare actual and predicted proportions, r^2 (weighted by the number of households in each cell) is calculated for the best linear fit between these measures. Also the slope and intercept of this best linear fit are calculated to reveal any systematic biases in the model, and the significance of the difference between this best linear fit and the 45-deg line can be estimated through use of an F-statistic.

To determine the sensitivity of this approach to different ways of defining cells, actual versus predicted comparisons were made for various cell formation criteria.

Figure 1. Predicted probability versus actual probability for 1-trip-maker case.



Specifically, for the 1-trip-maker model, 10 completely different sets of cells were formulated by specifying different sets of ranges for the independent variables. The average value for the best fit r^2's was $\bar{r}^2_{best} = 0.81$, and the variance was $\sigma_{r^2} = 0.016$. The relatively small variance in this statistic is an initial indication that the technique is rather insensitive to the manner in which the independent variables are divided in order to assign observations to cells as long as a degree of homogeneity is maintained within each cell and care is taken to create cells that have a similar number of observations.

The average slope and intercept of the best fit lines for the 10 sets of cells used are $\bar{\beta} = 0.84$ and $\bar{\alpha} = 0.076$ respectively; variances are $\sigma^2_\beta = 0.016$ and $\sigma^2_\alpha = 0.007$ respectively. These average slope and intercept values suggest that the model predicts too high for small actual probabilities and too low for large actual probabilities. Thus there appears to be a systematic conservative prediction bias in the 1-trip-maker model. The relatively small variance for both of these statistics is further evidence that the proposed technique is rather insensitive to empirical issues such as the range of the variables defining the cells.

For the 1-trip-maker case overall, the technique reveals a relatively good correspondence between actual and predicted proportions. However, further work is required to determine exactly how good this correspondence is in a statistical sense. Assessing the goodness of fit in the manner presented here begins to give one insights into how well the logit model predicts disaggregate behavior. That the proposed technique is applicable only where large samples are available should be emphasized. Developing this procedure for choice settings with more than 2 alternatives and in a more statistically rigorous sense is within the realm of important further research.

CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

Results from the initial tests of some hypotheses of the automobile-ownership theory are encouraging. The estimated coefficient values of the variables of the models derived from the theory are in all cases correctly signed and are significantly different from 0. Also traditional goodness-of-fit measures are at values that are acceptable for nonlinear estimation equations of the multinomial logit type. And a relatively good correspondence exists between predicted and actual probabilities for groups of hold-out observations in the 1-trip-maker segment model.

The current models developed from a specific theory of automobile-ownership decisions help to begin to identify causal mechanisms in urban-household travel behavior. However, much remains to be done before these models can be effectively applied in predicting automobile-ownership changes as results of transportation system changes. Sensitivity analyses are required for a number of variable definitions and assumptions of the models. In particular, the models need to be tested for different measures of destination attraction a_k, different accessibility cutoffs (used to define sets D_0 and D_1) and different urbanized area definitions. Improved understanding of how individual trip makers in the household interact in their uses of 1 or more family automobiles also must be sought. Finally, better understanding of the goodness of fit of probabilistic choice models, such as the multinomial logit model, must be obtained on both aggregate and individual observation levels before any of these types of models can be employed confidently in prediction. We hope that the goodness-of-fit technique proposed herein and the criticisms it encourages will lead toward a better understanding of the complex goodness-of-fit phenomena.

ACKNOWLEDGMENTS

We wish to thank Martin J. Beckmann, Technical University of Munich, and Richard L. Gustafson, formerly of General Motors Research Laboratories and currently an Oregon state representative, for their theoretical and methodological inputs. Appreciation also is extended to Larry L. Howson, formerly of General Motors Research Laboratories and currently with General Motors Transportation Systems Division, for his invaluable computer assistance and to Wilfred W. Recker and K. S. Krishnan, General Motors Research Laboratories, for their criticism.

REFERENCES

1. M. J. Beckmann, R. L. Gustafson, and T. F. Golob. Locational Factors in Automobile Ownership Decisions. Annals of Regional Science, Vol. 7, 1973, pp. 1-12.
2. M. J. Beckmann and T. F. Golob. Traveler Decisions and Traffic Flows. In Transportation and Traffic Theory (D. J. Buckly, ed.), A. H. and A. W. Reed, Sydney, Australia, 1974.
3. M. E. Beesley and J. F. Kain. Urban Form, Car Ownership and Public Policy: An Appraisal of Traffic in Towns. Urban Studies, Vol. 1, 1964, pp. 174-203.
4. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Massachusetts Institute of Technology, PhD dissertation, 1973.
5. D. Bloch and G. Watson. A Bayesian Study of the Multinomial Distribution. Annals of Mathematical Statistics, Vol. 38, 1967, pp. 1423-1435.
6. R. Bock. Estimating Multinomial Response Relations. In Contributions to Statistics and Probability (R. Bose, ed.), Univ. of North Carolina Press, Chapel Hill, 1969.
7. H. Botzow. An Empirical Method for Estimating Auto Commuting Costs. Highway Research Record 197, 1967, pp. 56-69.
8. D. Brand. Travel Demand Forecasting: Some Foundations and a Review. HRB Special Rept. 143, 1973, pp. 239-282.
9. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, March 1972.
10. N. Cherniack. Critique of Home-Interview Type O-D Surveys in Urban Areas. HRB Bulletin 253, 1960, pp. 166-188.
11. J. G. Cragg. Some Statistical Models for Limited Dependent Variables With Application to the Demand for Durable Goods. Department of Economics, Univ. of British Columbia, Discussion Paper 8, 1968.
12. J. G. Cragg and R. S. Uhler. The Demand for Automobiles. Department of Economics, Univ. of British Columbia, Discussion Paper 27, 1969.
13. E. Debreu. Representation of a Preference Ordering by a Numerical Function. In Decision Processes (R. M. Thrall, C. H. Coombs, and R. L. David, eds.), 1954.
14. H. D. Deutschman. Auto Ownership Revisited: A Review of Methods Used in Estimating and Distributing Auto Ownership. Highway Research Record 205, 1967, pp. 31-49.
15. R. T. Dunphy. Transit Accessibility as a Determinant of Automobile Ownership. Highway Research Record 472, 1973, pp. 63-71.
16. S. V. Ferrara. CATS Car Ownership Forecast: A New Approach. Chicago Area Transportation Study Research News, June 1965.
17. T. F. Golob. The Survey of User Choice of Alternate Transportation Modes. High Speed Ground Transportation Journal, Vol. 4, 1967, pp. 103-116.
18. T. F. Golob and M. J. Beckmann. A Utility Model for Travel Forecasting. Transportation Science, Vol. 4, 1971, pp. 87-92.
19. T. F. Golob, R. L. Gustafson, and M. J. Beckmann. An Economic Utility Theory Approach to Spatial Interaction. Papers of Regional Science Association, Vol. 30, 1973, pp. 159-182.
20. J. Gurland, I. Lee, and P. Dolan. Polychotomous Quantal Response in Biological Assay. Biometrics, Vol. 16, 1960, pp. 382-398.
21. D. T. Hartgen and G. H. Tanner. Individual Attitudes and Family Activities. High Speed Ground Transportation Journal, Vol. 4, 1970, pp. 439-467.
22. J. F. Kain and M. E. Beesley. Forecasting Car Ownership and Use. Urban Studies, Vol. 2, 1965, pp. 163-185.
23. C. F. Manski. Maximum Score Estimation of the Stochastic Utility Model of

Choice. School of Urban and Public Affairs, Carnegie-Mellon Univ., 1974.
24. D. McFadden. The Revealed Preferences of a Government Bureaucracy. Department of Economics, Univ. of California, Berkeley, 1968.
25. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1973.
26. W. Michelson. The Use of Social Statistics in Estimating Auto Ownership. Highway Research News 16, Dec. 1964, pp. 36-37.
27. J. H. Niedercorn and B. V. Bechdolt. An Economic Derivation of the Gravity Law of Spatial Interaction. Journal of Regional Science, Vol. 10, 1969, pp. 407-410.
28. W. Oi and P. Shuldiner. An Analysis of Urban Travel Demand. Northwestern Univ. Press, Evanston, Ill., 1962.
29. 1975 Transportation Plans. Penn-Jersey Transportation Study, Philadelphia, Vol. 3, May 1965.
30. Vehicle Registration Forecast. Pittsburgh Area Transportation Study, Technical Paper 3, June 1960.
31. 1985 Forecasts of Trucks and Passenger Vehicles Owned at Households. Puget Sound Regional Transportation Study, Seattle, Staff Rept. 9, April 1964.
32. P. R. Rassam, R. H. Ellis, and J. C. Bennett. The n-Dimensional Logit Model: Development and Application. Highway Research Record 369, 1971, pp. 135-147.
33. R. Shindler and M. G. Ferreri. Auto Ownership as Affected by Transportation System Alternatives. Traffic Engineering, Vol. 38, 1967, pp. 24-28.
34. P. R. Stopher and T. E. Lisco. Modelling Travel Demand: A Disaggregate Behavioral Approach—Issues and Applications. Proc., Transportation Research Forum, Vol. 11, 1969, pp. 195-214.
35. Growth, Change, and a Choice for 1990: Preliminary Plan, Southeast Michigan. Detroit Regional Transportation and Land Use Study, 1969.
36. H. Theil. A Multinomial Extension of the Linear Logit Model. International Economic Review, Vol. 10, 1969, pp. 251-259.
37. H. Theil. On the Estimation of Relationships Involving Qualitative Variables. American Journal of Sociology, Vol. 76, 1970, pp. 103-154.
38. H. Theil. Principles of Econometrics. John Wiley and Sons, Inc., New York, 1971.
39. States and Small Areas: Michigan. In U.S. Census of Housing: 1960, Vol. 1, Final Rept. HC(1)-24, U.S. Government Printing Office, 1962.
40. UTPS Reference Manual, August 1, 1974. Planning Methodology and Technical Support Division, Urban Mass Transportation Administration, Rept. UTP 40.74.1.5, 1974.
41. A. G. Wilson. A Statistical Theory of Spatial Distribution Models. Transportation Research, Vol. 1, 1967, pp. 253-270.
42. A. G. Wilson. Entropy in Urban and Regional Modelling. Pion Ltd., London, 1970.

DISCUSSION

Fred A. Reid, University of California, Berkeley

The paper by Burns, Golob, and Nicolaidis is a contribution to behavioral modeling and a utility theory of automobile ownership. Noteworthy points of the theory are expressing household choice in terms of the travel utilities of the individuals and maximizing budget and overall utility at the household level. Trip-frequency and destination-choice factors are included, although modal choice is considered only for adults in excess of the household automobile count. The theoretical development is valuable for identifying the construction of the utility function and the interplay of individuals in a household. However, if one is trying to develop accurate, manageable models in terms of available data, the theory seems too complex in relation to the number of parameters estimated (only 3 per population segment) and the restrictive assumptions necessary for calibration. The relation between trip level of service and attraction variables for all types of trips is completely specified by the theory rather than by allowing separate parameters to be estimated or different functional forms for alternate theories to be tested. For example, service variables have been found to best describe travel behavior entering as linear rather than logarithmic differentials in other studies.

The failure to characterize the individual attributes of the household work trip is a weakness when one considers the role these major trips play in ownership and because of the generality of the destination-attraction variable used. As only the percentage of total population and employment at a destination, this variable also does not distinguish non-work-trip ends as being unique to a household.

The assumption that owned automobiles are always used.if they are available seems particularly restrictive. It leaves modal-choice alternatives to only a portion of the household members, weakening the advantage of the individual utility maximization.

The paper has shown some useful theoretical development, but I believe it has raised more questions than it has answered.

AUTHORS' CLOSURE

We are pleased that Reid viewed this paper as a contribution to behavioral travel-demand modeling in general and to the understanding of automobile demand specifically. We thank him for his comments. Comments on his discussion are warranted.

We are aware of the limitations of a rigid theoretical approach. However, we judge that attempting to assess causal mechanisms in urban household automobile-ownership decisions by merely studying ad hoc empirical correlations is a much less satisfactory alternative. Consequently, the development of an economic theory of decision-making behavior placed significant restrictions on the empirical study performed to test the hypotheses of the theory. The objective of establishing a valid causal model with sound theoretical underpinnings justified acceptance of these restrictions. Given that a predefined theory was rigidly followed, greater value can be placed on the significance of the empirical results because probabilities of uncovering random phenomena are minimized.

Admittedly, the existing theory assumes independence among household members in their trip-making behavior. As a result, the utility of additional automobiles to a household is not accurately modeled. Attempts to account for the interdependence of household members' trip-making behavior significantly complicates the model. These complications are not warranted in an initial study of automobile-ownership decisions. As noted in the original paper, however, such concerns, which were uncovered by following the theoretical approach, indicate important directions for further research.

The theoretical model can be adapted to include

1. Purpose-specific attraction measures;
2. Variances in the cost of automobile ownership due to residential location, household income, and type of automobile (first, second, or third household automobile); and
3. Travel times broken down into walking, waiting, and in-vehicle time components.

Including these features in the initial empirical study was prohibited by data availability. Also the final model allows assessment of impacts of improvements in public transportation or changes in automobile travel characteristics on automobile ownership. Improvements in public transportation are reflected in reduced travel times and increased service areas. Such changes can be incorporated directly into the model transit utility term, making assessment of the effects of public transportation improvements a relatively straightforward procedure. In a similar manner, the effects of increased or decreased automobile travel times can easily be reflected in the automobile travel utility terms.

We would like to emphasize again, as we did in our conclusion, that the objective of our research was to identify the structure of causal mechanisms in urban-household travel behavior. We are aware that, before the models resulting from this research can be effectively applied in predicting changes in automobile-ownership levels as results of transportation system changes, sensitivity analyses will be required to test further variable definitions and assumptions of the model. Specific areas of future research were covered in the paper.

TIME-STABILITY ANALYSIS OF TRIP-GENERATION AND PREDISTRIBUTION MODAL-CHOICE MODELS

Robert L. Smith, Jr.,* University of Wisconsin—Madison; and
Donald E. Cleveland, University of Michigan

Data from home-interview surveys in Detroit for 1953 and 1965 were used to test the time stability of disaggregate trip-generation and predistribution modal-choice models. Initial cross-classification analysis showed 4 to 18 percent increases in household trip-generation rates for households with cars available. A statistical test of the overall time stability of multiple linear regression trip-generation equations indicated that the equations were not stable unless non-trip-making households were removed. The individual regression coefficients also were tested for time stability, and, despite the lack of overall statistical time stability, disaggregate work- and home-based trip-generation equations for 1953 produced reasonable estimates of 1965 zone-level trips. Disaggregate regression equations for the automobile-driver and bus modes also were found to exhibit only limited time stability. Interaction between cars available and number of persons employed was particularly important in explaining bus trips. Tests of the time-stability assumption at the zone level were limited by the lack of zone-level interaction variables.

THE CONVENTIONAL sequential models of urban travel demand (UTD models) require 3 basic assumptions for use in forecasting: (a) independent variables can be accurately forecast; (b) models provide an accurate, behaviorally correct simulation of base-year travel demand; and (c) model variables, structure, and parameters are stable over time (1). Early researchers in transportation planning were well aware of the need for accurate forecasts of independent variables. During the 1960s, considerable effort was devoted to developing sophisticated urban land use activity models to provide the required forecasts. Although the problem of producing accurate forecasts has proved more intractable than initially thought, the models can generate alternative land use patterns ranging from trend to normative statements of future development patterns such as centralization, radial corridors, or satellite development (2). The last 2 assumptions required for forecasting travel demand are closely related. Behavioral models that accurately predict base-year travel also should be valid in some future year. Deutschman (1), however, has argued that trip-generation models that produce a good fit for the base year nevertheless may fail completely when used for forecasting. Clearly, the goodness of fit of base-year data should not be the only criterion for model selection. Behavioral models are needed to provide not only time stability but also meaningful responses to changes in transportation systems, land use activity patterns, and socioeconomic conditions.

Considerable research has been devoted to developing better behavioral UTD models. Emphasis has focused on modal-choice models and, to a lesser extent, trip-generation models. Reported research, however, largely has ignored the other possible sources of error in forecasting travel demand. The attitude toward the accuracy of forecasts of independent variables appears to be that, if an independent variable is behaviorally

*Mr. Smith was with the Southeast Michigan Council of Governments when this research was performed.

significant, then some attempt should be made to forecast it. Roberts (3), however, has argued that, in view of the current difficulties in making accurate long-range forecasts of independent variables, the current UTD models are valid only for the short range. Such a view could also make time-stability errors relatively less important. If, however, the short range includes time periods of 5 to 10 years (which are not unreasonable periods for implementation of major transportation projects), time stability still may be an important criterion. Time-stability analyses have been limited by the lack of adequate time-series data. Even where data have been available for 2 or more time periods, operational pressures to produce travel demand forecasts as quickly and cheaply as possible generally have resulted in the use of only the latest data for model calibration and forecasting.

PURPOSE

The primary purpose of this research was to reevaluate the Detroit trip-generation model so that any new data requirements could be incorporated in the development of a new regional activity allocation model (4). The new data requirements also would provide a framework for an updating of the 1965 home-interview survey data by using 1970 census data.

In reevaluating the Detroit trip-generation model, emphasis was placed on developing a behavioral trip-generation model that would exhibit time stability. Because relatively consistent home-interview survey data were available for both 1953 and 1965, disaggregate trip-generation models were developed for both years and then were compared graphically and statistically for stability. Predistribution modal-choice models for both years also were developed to aid in understanding how changes in modal choice affect trip generation.

The substantial changes in population, residential density, and job location and in automobile ownership, household income, and level of service of the transportation system between 1953 and 1965 provided a significant test of trip-generation-model time stability. The sharp decline in transit use from 1953 to 1965 (16.2 percent of person trips by bus in 1953 versus 4.8 percent by bus in 1965) subjected the predistribution modal-choice model to an even more severe test of time stability.

RESEARCH METHODOLOGY

Household files with merged trip-record data for 1953 and 1965 were developed to be as nearly compatible as possible. The major limitation of the 1953 file was the lack of individual household income data. The 1953 variable, cars owned, was assumed to be equivalent to the 1965 variable, cars available; a compatible stage of life cycle variable was developed as a function of family size, age of youngest trip maker, and age of the head of the household. An additional limitation of the 1953 household file was the lack of complete data for variables obtained from the trip files. If no trips were made, then no data were available for these variables. Thus use of the trip file variables was limited to models for trip-making households.

A summary of the major independent and dependent variables that were compatible for 1953 and 1965 is given in Table 1. All but 1 of the variables that generally have been considered to be most significant in household-level trip-generation analysis are included (5, p. 96). The variable omitted is number of persons 16 years old or older who drive. One variable, AREA, requires some explanation. The AREA variable stratifies the region into 4 roughly concentric rings centered on the central business district (CBD): inner city, city center, suburbs, and rural area. The rural area is outside the 1953 study-area boundary; thus, no data are presented for this location. The AREA variable is a substitute for such variables as density, distance from CBD, and, to some extent, accessibility.

Recently, there has been considerable discussion of the desirability of including measures of the level of service provided by the transportation system at each stage of the urban transportation planning modeling process (6, 7). However, little empirical evidence exists to support the notion that transportation accessibility as currently measured is significant in explaining trip generation. For a household-level model Kannel (8, p. 116) concluded that "the effect of accessibility [to employment] on trip production rates would appear to be an indirect effect due to its influence on auto ownership." For zone-level models, both Nakkash (9) and Gur (10) found that accessibility variables contributed little to the explanatory power of the models. Nakkash (9) found that a simple stratification of zones into central and noncentral areas was more significant than inclusion of accessibility variables. The number of work trips is relatively inelastic with respect to accessibility or transportation cost. The level of non-work-trip making appears to be primarily a function of the number of automobiles available although the level of automobile ownership may be affected by transportation system accessibility, as shown by Kannel (8). Dunphy (11) has shown that automobile ownership also may be affected by transit accessibility to employment, thus holding 2 variables, income and family size, constant.

Multiple linear regression analysis was selected to develop trip-generation relationships for 1953 and 1965 because of the ease with which statistical measures of both goodness of fit and time stability can be obtained. Cross-classification techniques were used to examine the extent to which the data met the standard assumptions required for regression analysis. In addition, the automatic interaction detection program (AID) was used to identify potential interaction terms (12).

Four different approaches to trip-generation time-stability evaluation were used:

1. Graphical comparison,
2. Test of overall equality of regression-equation coefficients,
3. Test of equality of individual regression coefficients, and
4. Prediction of 1965 zone and district trips by using the 1953 equations.

The second approach used Chow's test for the equality of 2 sets of linear regression coefficients (13). In Chow's test of the equality of 2 sets of regression coefficients, the null hypothesis of equality of the regression coefficients for the 2 years (H_0: $\beta_1 = \beta_2$) is rejected at a $(1 - \alpha)$ percent level of confidence if the test statistic F is greater than $F_{1-\alpha}$ with k and $(m + n - 2k)$ degrees of freedom. F is computed

$$F = \frac{(Q_1 - Q_2)/k}{Q_2/(m + n - 2k)} \tag{1}$$

where

Q_1 = sum of squared errors from pooling the observations,
Q_2 = sum of squared errors from separate regressions for the 2 years,
m = number of observations in year 1,
n = number of observations in year 2, and
k = number of independent variables plus 1.

The practical application of the technique requires both a separate regression for each year and a regression on the pooled observations for both years. The difference between Q_1 and Q_2 provides a measure of the closeness of the 2 sets of regression coefficients. If the regression equations are identical, the difference between Q_1 and Q_2 will be 0.

The third approach used the time interval as a dummy variable to test each regression coefficient for change over time. The time-period dummy variable T is included as an interaction term with every independent variable, including the constant term. For example, the regression equation

$$HB = a_0 + a_1CARA + a_2NRES \tag{2}$$

becomes

$$HB = a_0 + b_0T + (a_1 + b_1T)CARA + (a_2 + b_2T)NRES \tag{3}$$

where each interaction term has been combined with its respective independent variable. The coefficient of each independent variable is tested for time stability under the null hypothesis that the interaction term regression coefficient b_i equals 0 (H_0: $b_i = 0$).

Identical analyses were used to develop predistribution modal-choice models and evaluate their time stability. Considerable attention was given to the development of appropriate interaction terms.

TRIP-GENERATION MODEL DEVELOPMENT

Primary emphasis was placed on developing trip-generation equations for TTF, HB, and WK. Equations also were developed for PB, SR, and SHOP trip purposes.

The 2 types of independent variables found to be most important in explaining total trip generation were (a) some measure of household size and (b) some measure of household economic status. The available household-size variables for the Detroit area were LC, NRES, and FIVE. CARA was the only economic variable available for both years. The impact of INC, however, could be analyzed at the household level for 1965. EMP was important for explaining work trips. The other available variables given in Table 1 also were evaluated for significance in explaining the various trip purposes.

An initial cross-classification analysis provided an indication of the extent to which the standard assumptions required for regression analysis were met as well as a graphical measure of the degree of time stability. The cross-classification of TTF by CARA and NRES for both 1953 and 1965 (Figure 1) showed relatively consistent change over time. The mean daily trip rates for household-size classes with 1 or more cars available were approximately 10 percent greater in 1965 than in 1953. (The actual range was 4 to 18 percent; the increase for the individual car-available classes ranged from 1 to 9 percent.) The relatively uniform upward shift suggests an additional income effect or a uniform regional increase in accessibility. A shift from neighborhood walk trips (no data available) to vehicle trips for shopping, social-recreational, and similar purposes also might have contributed to the increase.

In contrast, households in each household-size class with no cars available made approximately 20 percent fewer trips in 1965 than in 1953. The no-car-available class in 1965 was composed primarily of poor and elderly people, many of whom reported not making any trips. Underreporting of trips also might be more of a problem here.

When non-trip-making households were removed, the difference in trip making between 1953 and 1965 for the household-size classes with no cars available was reduced from a 20 percent decline to a range of -3.9 to +8.7 percent change. Almost all households with 1 or more cars available made trips. Thus little change in the trip rates of the 1- and 2-cars-available classes as a function of NRES occurred when non-trip-making households were removed. Further stratification by AREA of the CARA versus NRES curves for trip-making households showed a generally higher level of suburban (AREA = 3) trip making and a lower level of inner city (AREA = 1) trip making in 1965. Trip making by city center residents (AREA = 2) remained relatively constant. Figure 2 shows the relationships for a family of 4 (NRES = 4).

Cross tabulation of home-based WK for both years by EMP and CARA showed the expected essentially constant rate of WK/person employed. To expect WK generation to be independent of INC and CARA is reasonable. There should also be little change over time unless, for example, the 4-day work week were to be adopted widely.

Table 1. Household variables.

Variables	Designation
Independent	
Cars available, 1965	CARA
Cars owned, 1953[a]	CARA
Income[b]	INC
Stage in the family life cycle	LC
Number of household residents[a]	NRES
Number of persons 5 years old and older[a]	FIVE
Number of persons employed	EMP
Race of household head	RACE
Type of structure (single or multiple)[a]	STR
Occupation of household head	OCC
Sex of household head	SEX
Labor force status of household head	LF
Age of youngest trip maker	Y
Location in the region[a]	A1, A2, A3, A4
Small area location[a]	ZONE
Dependent	
Total factored person trips	TTF
Total home-based person trips	HB
Work trips	WK
Personal business trips	PB
Social and recreational trips	SR
Shopping trips	SHOP

[a]Available for all households in 1953.
[b]Available only at the census tract level for 1953.

Figure 1. Household trip rates.

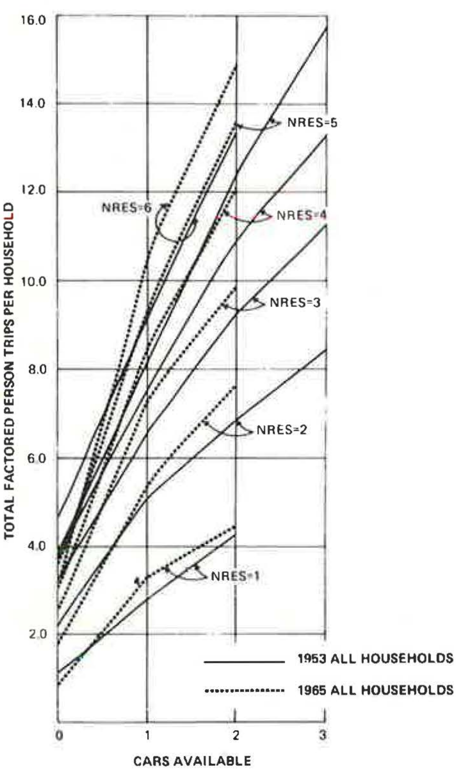
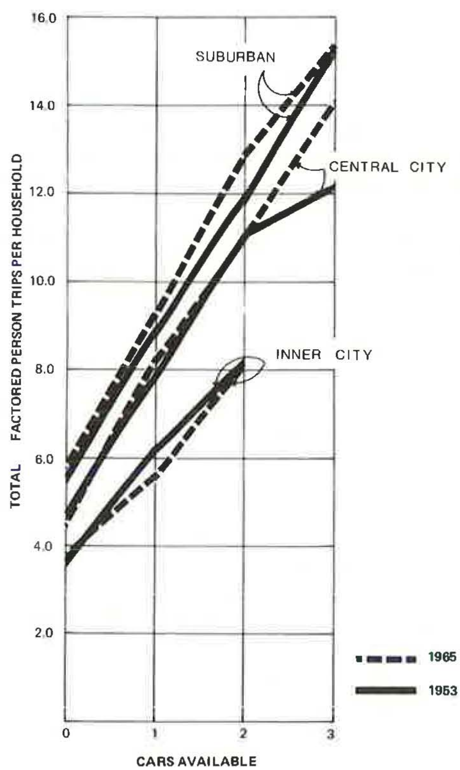


Figure 2. Trip rates for trip-making households with 4 residents.



The linearity assumption required by regression analysis (unless dummy variables are used) appears to be generally satisfied for the TTF trip relationships so that little advantage would be obtained from using dummy variables for CARA or NRES. The main restriction on the regression analysis appeared to be the lack of normal error terms with constant variance. The distributions of the dependent variables were skewed, which indicates that the distributions of the error terms probably were not normal. Also the variance of the dependent variables tended to increase with the independent variables, which indicates heteroscedastic error terms. Similar results have been reported by Oi and Shuldiner (14) and Kannel (8). Thus the statistical reliability of the significance tests on the regression coefficients is likely to be overstated.

The graphical cross-classification analysis indicated that the additivity assumption required for regression analysis generally was satisfied. There was little evidence of interaction among the independent variables. More detailed analysis of interaction by using the AID program confirmed the assumption of negligible interaction.

Household-level trip-generation regression equations were developed for both years by using a systematic 1-in-8 sample of the more than 40,000 households in both the 1953 and 1965 household files to reduce computing costs (Table 2). The analysis was concentrated on the home-based-trip purpose because home-based trips provided the control total for the individual home-based-trip purposes in the 1965 Detroit trip-generation model. Selecting a household-size variable for the HB-FIVE equation resulted in a slightly higher coefficient of determination r^2 than for the HB-NRES equation. However, for forecasting purposes NRES probably would be better because it is available in the 1970 census transportation planning package. LC provided greater explanatory power than did either NRES or FIVE. However, even if LC were available for all households, the potential error in forecasting it probably would outweigh the improvement in base-year accuracy. The AREA variable was significant for the home-based equation but not for the work or personal-business equations.

The trip-generation equations for trip-making households only were developed by using the same independent variables as were used in the all-households equations. The coefficients of NRES were essentially the same as for the comparable all-households equation; however, the constant terms and the coefficients of CARA changed substantially. The trip-making-only equations exhibited consistently lower coefficients of determination than did the all-households equations.

TIME-STABILITY EVALUATION

Test of Overall Stability

The results of the test of the null hypothesis of no difference in the trip-generation equation regression coefficients (H_0: $\beta_{53} = \beta_{65}$) are given in Table 3. Only 2 equations were concluded to be stable at the 1 percent level of significance—the HB equation for trip-making households only with CARA and NRES as independent variables and the SR equation for all households with CARA, NRES, and AREA as independent variables. Even these 2 equations were not stable at the 5 percent level of significance. Thus the statistical analysis indicates that, in general, the overall differences in trip-generation rates that were observed both in the graphical analyses and in the comparison of the individual regression equations for the 2 years are statistically significant.

Test of Regression Coefficient Stability

The test of the time stability of the individual regression coefficients confirmed the results of the overall time-stability analysis for the home-based purpose (Table 4). The constant term was found to change over time (at the 1 percent level) for all of the HB equations except the trip-making-households-only equation with CARA and NRES as independent variables. The results for the individual home-based-trip purposes (WK, PB, SR, and SHOP), however, showed that stability of each of the coefficients in an trips are available from government or private data sources.

2. A trip-distribution model for small- and medium-sized urban areas can be calibrated by using available trip-end information.

3. The necessary information for traffic planning, the average daily traffic and the peak-hour volumes, can be deduced from the home-based work trips.

GRAVITY-DISTRIBUTION MODEL FOR SMALL- AND MEDIUM-SIZED URBAN AREAS

Following Evans (3) and Sasaki (9), one can write a 2-way, constrained gravity model as

$$T_{ij} = r_i s_j F(C_{ij}) \qquad (1)$$

where

T_{ij} = number of trips originated from zone i and destined to zone j as predicted by a 2-way gravity model;
$F(C_{ij})$ = distribution (impedance) function, which is a function of the travel cost C_{ij}; traditionally, travel cost is expressed in minutes of travel time; and
r_i and s_j = normalization factors established so that trip productions and trip attractions predicted by the gravity model become equal to the original trip productions and attractions (definition of a 2-way, constrained gravity model).

Therefore, r_i and s_j are solutions of the following equations:

$$\sum_j r_i s_j F(C_{ij}) = P_i \qquad (2)$$

$$\sum_i r_i s_j F(C_{ij}) = A_j \qquad (3)$$

which also satisfy

$$\sum_i \sum_j r_i s_j F(C_{ij}) = T \qquad (4)$$

where

P_i = trip production of zone i,
A_j = trip attraction of zone j, and
T = total trip exchange within the system.

This notation of the gravity model is basically the same as the conventional use of the 2-way, constrained gravity model (3). In the conventional use, trip attractions are iteratively changed for meeting the trip end constraints; in this notation, the normalization factors r_i and s_j are iteratively set to meet these constraints. This notation is being used in this paper only for clarification. The distribution function quantified for equation 1 can be used for the conventional use of the gravity model without any change.

Table 2. Trip-generation equations.

Households	Year	Regression Equation[a]	r^2	Standard Error	Mean
All[b]	1965	HB = -0.65 + 2.43CARA + 0.93NRES	0.355	3.93	5.61
	1953	HB = 0.11 + 2.02CARA + 0.81NRES	0.276	3.42	4.72
	1965	HB = -0.84 + 2.30CARA + 1.14FIVE	0.377	3.86	5.61
	1953	HB = -0.29 + 1.82CARA + 1.13FIVE	0.314	3.33	4.72
	1965	HB = -1.41 + 2.20CARA + 0.93NRES + 0.76A2 + 1.48A3	0.364	3.90	5.61
	1953	HB = -0.31 + 1.87CARA + 0.80NRES + 0.67A2 + 0.89A3	0.282	3.41	4.72
	1965	WK = +0.08 + 1.63EMP	0.582	1.01	1.82
	1953	WK = -0.06 + 1.65EMP	0.582	1.05	2.01
	1965	PB = -0.18 + 0.62CARA + 0.18NRES	0.102	1.97	1.21
	1953	PB = -0.07 + 0.46CARA + 0.11NRES	0.069	1.43	0.72
Trip making only[c]	1965	HB = +0.17 + 2.10CARA + 0.92NRES	0.272	4.02	6.40
	1953	HB = +0.90 + 1.84CARA + 0.76NRES	0.233	3.38	5.39
	1965	HB = -0.81 + 1.87CARA + 0.92NRES + 1.01A2 + 1.68A3	0.283	3.99	6.40
	1953	HB = +0.47 + 1.69CARA + 0.75NRES + 0.63A2 + 0.91A3	0.240	3.36	5.39
	1965	WK = +0.20 + 1.59EMP	0.534	1.03	2.08
	1953	WK = -0.04 + 1.66EMP	0.501	1.09	2.30

[a]All coefficients are significant at the 1 percent level.
[b]Sample size: 2,586 for 1965 and 2,529 for 1953.
[c]Sample size: 2,265 for 1965 and 2,216 for 1953.

Table 3. Summary of overall time-stability test, $H_0 : \beta_1 = \beta_2$.

Dependent Variable	Independent Variables	r^2 1965	r^2 1953	Degrees of Freedom	F (ratio of mean standard error)
HB	CARA, NRES	0.355	0.276	3	5.22
HB	CARA, NRES, A2, A3	0.364	0.282	5	4.68
HB	CARA, FIVE, A2, A3	0.388	0.322	5	4.20
HB[a]	CARA, NRES	0.272	0.233	3	3.65[b]
HB[a]	CARA, NRES, A2, A3	0.283	0.240	6	3.51
WK[a]	EMP	0.534	0.501	2	11.9
WK	EMP	0.582	0.582	2	8.36
PB	CARA, NRES, A2, A3	0.108	0.070	5	10.46
SR	CARA, NRES, A2, A3	0.115	0.085	5	2.83[b]
SHOP	CARA, NRES, A2, A3	0.110	0.082	5	5.99

[a]Includes trip-making households only.
[b]Stable at 1 percent level.

Table 4. Summary of regression slope and intercept stability.

Households	Regression Equation
All	HB = 0.11 - 0.76T + (2.03 + 0.41T)CARA + (0.81 + 0.12T)NRES t values = 3.03[a] 17.3 2.67[a] 16.2 1.80
All	HB = -0.30 - 1.11T + (1.87 + 0.33T)CARA + (0.80 + 0.13T)NRES + (0.67 + 0.09T)A2 + (0.89 + 0.59T)A3 t values = 3.45[a] 15.4 2.06 16.0 1.96 3.35 0.27 4.36 1.84
All	HB = -0.75 - 0.92T = (1.64 + 0.39T)CARA + (1.12 + 0.02T)FIVE + (0.69 + 0.11T)A2 + (0.98 + 0.61T)A3 t values = 2.93[a] 13.6 2.52 19.8 0.27 3.51 0.34 4.94 1.97
Trip making only	HB = 0.92 - 0.75T + (1.84 + 0.26T)CARA + (0.75 + 0.16T)NRES t values = 2.52 15.4 1.56 14.1 2.30
Trip making only	HB = 0.50 - 1.31T + (1.69 + 0.18T)CARA + (0.74 + 0.17T)NRES + (0.62 + 0.39T)A2 + (0.91 + 0.78T)A3 t values = 3.46[a] 12.8 1.04 13.9 2.46 2.80 1.06 4.02 2.18
All	WK = -0.06 + 0.14T + (1.65 - 0.02T)EMP t values = 2.67[a] 60.6 0.56
All	PB = -0.07 - 0.11T + (0.46 + 0.16T)CARA + (0.11 + 0.07T)NRES t values = 0.97 8.33 2.27 4.54 2.38
All	SR = -0.35 - 0.24T + (0.44 - 0.10T)CARA + (0.18 + 0.11T)NRES + (0.36 - 0.15T)A2 + (0.45 + 0.06T)A3 t values = 1.47 7.01 1.22 6.96 3.19[a] 3.42 0.90 4.29 0.35
All	SHOP = -0.24 - 0.17T + (0.28 + 0.07T)CARA + (0.12 + 0.04T)NRES + (0.21 + 0.12T)A2 + (0.46 + 0.21T)A3 t values = 1.33 5.67 1.05 5.94 1.57 2.58 0.92 5.61 1.67

[a]Significant at the 1 percent level and therefore unstable.

equation does not guarantee the overall time stability of the equation. Neither the PB nor the SHOP equation had any unstable coefficients, yet neither equation exhibited overall time stability. In contrast, the SR equation exhibited overall time stability, but the coefficient of the independent variable NRES changed over time. The results of this test should not be overemphasized because the assumptions required for multiple linear regression analysis were not analyzed for the individual home-based-trip purposes.

Application of the Zone Level

The final test of the time stability of the regression equations was to forecast 1965 zone trips by using the equations developed for 1953. The 1953 HB trip-generation equation with CARA, NRES, and AREA as independent variables was applied to 1965 zone-level data. The resulting estimates of 1965 zone HB trips were reasonable (r^2 for actual versus estimate was 0.950). The 1965 zone estimates produced by the comparable 1965 HB equation were only slightly more accurate than those produced by the 1953 HB equation (Figure 3). When stratified by AREA, the 1953 HB equation (and probably the 1965 HB equation as well) was more accurate for estimating the city center and suburban areas than it was for the inner city and rural areas, which indicates that additional study of the latter 2 areas is needed.

Comparison With Other Urban Areas

Ideally, household-level trip-generation regression equations developed in one urban area also should be valid for other urban areas. Comparison of home-based trip-generation equations for urban areas ranging in size from 250,000 (Madison, Wisconsin) to more than 14,000,000 (New York City) shows substantial variations in the regression coefficients although the same independent variables are significant for all of the U.S. cities (Table 5). Part of the variation may be attributed to differences in data collection and definitions of the variables. The wide variation in the CARA coefficients also may be the result of differences in transit service and household income. There is relatively little variation in the FIVE coefficient. Madison, Wisconsin, may be a special case because of the large college-student population.

A comparison of household-level regression equation slope and intercept stability between Pittsburgh and Detroit shows generally similar results although the magnitude of individual regression coefficients differs substantially (Table 6). The slopes for the SHOP equation are stable (at the 1 percent level), and slopes for the TTF equation are generally unstable. The independent variables for other purposes in the Pittsburgh study were not compatible with the Detroit variables.

PREDISTRIBUTION MODAL-CHOICE ANALYSIS

The 2 most important trip purposes, home based and work, were analyzed for both the automobile-driver and bus modes. In contrast to the person-trip-generation relationships, the AID analysis indicated significant interactions between the 2 primary independent variables for the bus mode—CARA and EMP. The subsequent cross-classification analysis of the bus purposes indicated that the interaction variable, EC, which is defined as the nonnegative difference between the number of employees in the home and the number of cars available, that is,

$$\mathrm{EC} = \begin{cases} (\mathrm{EMP} - \mathrm{CARA}) \geq 0 \\ 0 \quad \text{otherwise} \end{cases}$$

should be a good predictor of bus-work and bus-home-based trips. A graphical analysis of the stability of the bus-work trip relationships is shown in Figure 4. As expected, a generally lower level of transit use is observed in 1965.

Statistical analysis of the bus-work and bus-home-based regression equations for trip-making households indicated that none of several sets of independent variables provided overall time stability. When only households with convenient access to transit were considered (households in the bus service area), however, stable relationships for both the work and the home-based bus purposes were found with EC and AREA as dummy variables. The individual regression coefficients also were found to be stable for both equations. Additional research is needed to evaluate the feasibility for forecasting EC. EC is likely to be a complex function of the income level and the level of transit service.

As for the bus purposes, the AID analysis for the automobile-driver purposes indicated interaction between CARA and EMP. Subsequent cross-classification analysis indicated that the interaction variable, CE, defined as the nonnegative difference between CARA and EMP

$$\mathrm{CE} = \begin{cases} (\mathrm{CARA} - \mathrm{EMP}) \geq 0 \\ 0 \quad \text{otherwise} \end{cases}$$

should be a good predictor of automobile-driver trips. CE was significant as an explanatory variable for both automobile-driver-WK and automobile-driver-HB regression equations; however, neither equation exhibited overall stability for the entire region. Within the bus service area the automobile-driver-HB equation with CARA, FIVE, and AREA as independent variables exhibited overall stability. The individual regression coefficients also were stable.

Application of the bus-purpose equations to estimate zone bus trips was limited by the lack of interaction variables at the zone level. Reasonable estimates of zone trips for the 2 automobile-driver purposes, however, were obtained. The error curves (cumulative percent of zones versus percentage of error in the zone estimate) were virtually identical to the error curves for person-trip purposes, which is reasonable when one considers the low level of transit use in Detroit.

CONCLUSIONS

Cross-classification analysis of 1953 and 1965 total person-trip generation as a function of cars available and number of persons in the household showed unexplained differences in the trip-generation rates of 10 to 20 percent. The largest differences were observed in the no-car-available class. Thus, when non-trip-making households (which were concentrated in the no-car-available class) were removed, the resulting regression equation was concluded to be stable.

The relatively uniform increase in trip rates for all automobile-owning household-size classes probably is due to changes in income, regional accessibility, and the level of walking trips. Additional time-series data are needed to evaluate the impact of these variables. In the absence of such time-series data, the disaggregate trip-generation relationships for 1965 and 1953, as shown by the zone-level estimates, can provide an upper and a lower bound for reasonable estimates of future trip generation in Detroit.

The high degree of explanatory power of the work-trip-generation equation suggests that peak-hour UTD models based on work trips should be developed for Detroit. The lack of time stability probably was due primarily to differences in defining the employment variable between 1953 and 1965 rather than any inherent change in the level of work-trip making. Peak-hour models would be particularly useful because data are available from the 1970 census transportation planning package on work-trip and employment patterns. Thus the updating of the 1965 travel survey data is provided for.

The predistribution modal-choice analysis indicates the importance of the joint consideration of the number of persons employed and the number of cars available at the

Figure 3. Zone test of time stability.

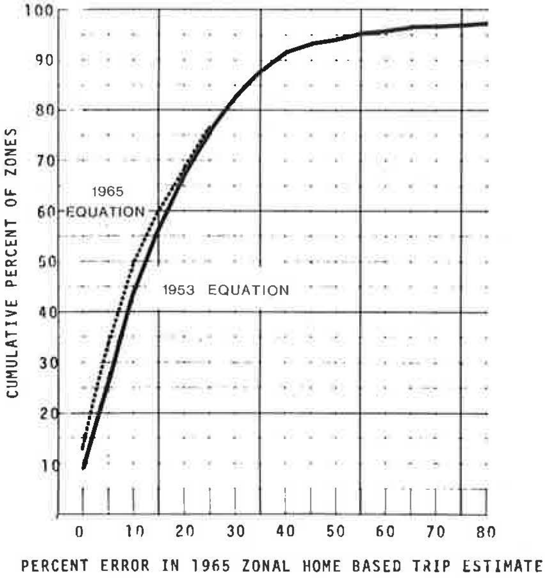


Table 5. Comparison of household-level home-based trip-generation equations for several urban areas.

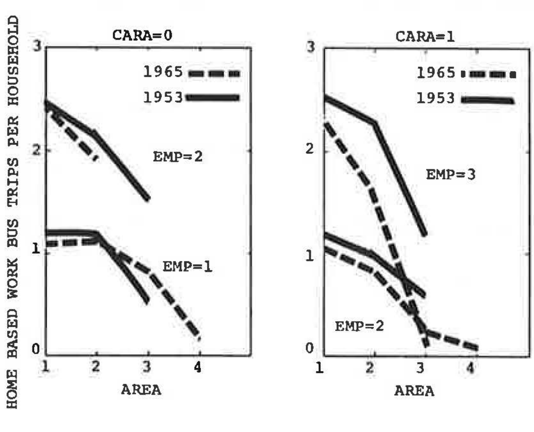
Urban Area	Year	Regression Equation	r^2	Standard Error	Mean
Madison, Wisconsin (5)	1962	HB = 0.69 + 1.94CARA + 1.39FIVE	0.36	3.89	5.20
Glamorgan, Wales (15)	—	HB = 0.91 + 1.07CARA + 1.07EMP	0.384	2.33	—
New York City (16)	—	HB = 0.24 + 3.17CARA + 1.06FIVE	0.309	5.03	5.87
Indianapolis (8)	1964	HB = -0.19 + 3.17CARA + 1.15FIVE	—	—	—
Detroit	1965	HB = -0.84 + 2.30CARA + 1.14FIVE	0.377	3.86	5.61
Detroit	1953	HB = -0.29 + 1.82CARA + 1.13FIVE	0.314	3.33	4.72

Table 6. Comparison of trip-generation slope and intercept time stability for Pittsburgh and Detroit.

Urban Area[a]	Regression Equation	r^2	Standard Error	Mean
Pittsburgh (17)	SHOP = 0.18 + 0.51T + (0.33 - 0.27T)CARA + (0.08 - 0.01T)NRES	0.030	—	—
	t values = 2.63[b] 2.11 0.11			
Detroit	SHOP = -0.07 + 0.0T + (0.36 + 0.10T)CARA + (0.13 + 0.04T)NRES	0.097	1.50	0.88
	t values = — 7.51 1.60 6.26 1.37			
Pittsburgh (17)	TTF = 1.66 + 1.82T + (2.00 - 0.38T)CARA + (0.57 - 0.69T)NRES	0.208	—	—
	t values = 3.44[b] 1.10 2.81[b]			
Detroit	TTF = 0.16 - 1.26T + (3.13 + 80T)CARA + (0.98 + 0.29T)NRES	0.289	6.05	7.33
	t values = 3.06[b] 16.3 3.20[b] 12.0 2.69[b]			

[a]Pittsburgh data are for 1958 and 1967. Detroit data are for 1953 and 1965.
[b]Significant at the 1 percent level.

Figure 4. Bus-work-trip time stability for households with no car and 1 car available.



household level in explaining transit use, particularly for work trips. In Detroit, significant transit use generally occurs only for households in which the number employed exceeds the number of cars available. Thus short-term transit service improvements are not likely to attract workers who have a car available. Over a longer time period, however, the introduction of competitive transit service may result in a decision not to replace the second car or not to buy a car when a family member joins the labor force.

ACKNOWLEDGMENT

We would like to thank the Southeast Michigan Council of Governments for making this research possible. Much of the research reported here was conducted as part of Robert L. Smith's doctoral research program in civil engineering at the University of Michigan (18).

REFERENCES

1. H. D. Deutschman. Establishing a Statistical Criterion for Selecting Trip Generation Procedures. Highway Research Record 191, 1967, pp. 39-52.
2. Urban Development Models. HRB Special Rept. 97, 1968.
3. P. O. Roberts. Demand Forecasting for Long-Range and Contemporary Options. HRB Special Rept. 143, 1973, pp. 52-71.
4. R. L. Smith. Time-Stability Analysis of the Trip Generation and Pre-Distribution Mode-Choice Models. Southeast Michigan Council of Governments, Detroit, Sept. 1973.
5. Guidelines for Trip Generation Analysis. Federal Highway Administration, 1967.
6. D. Brand. Theory and Method in Land Use and Travel Forecasting. Highway Research Record 422, 1973, pp. 10-20.
7. M. L. Manheim. Practical Implications of Some Fundamental Properties of Travel-Demand Models. Highway Research Record 422, 1973, pp. 21-38.
8. E. J. Kannel. A Disaggregate Analysis of Urban Travel Behavior. Purdue Univ., joint highway research project, July 1972.
9. T. Z. Nakkash. Activity-Accessibility Models of Trip Generation. Purdue Univ., PhD dissertation, 1969.
10. Y. Gur. An Accessibility Sensitive Trip Generation Model. Chicago Area Transportation Study, Oct. 1971.
11. R. T. Dunphy. Transit Accessibility as a Determinant of Automobile Ownership. Highway Research Record 472, 1973, pp. 63-71.
12. J. A. Sonquist. Multivariate Model Building: The Validation of a Search Strategy. Institute for Social Research, Univ. of Michigan, Ann Arbor, March 1970.
13. G. C. Chow. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica, Vol. 28, July 1960, pp. 591-605.
14. W. Y. Oi and P. W. Shuldiner. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, Ill., 1962.
15. A. A. Douglas and R. J. Lewis. Trip Generation Techniques, Part 4. Traffic Engineering and Control, Feb. 1971.
16. H. Kassoff and H. O. Deutschman. Trip Generation: A Critical Appraisal. Highway Research Record 297, 1969, pp. 15-30.
17. N. Ashford and F. M. Holloway. The Performance of Trip Generation Equations. Transportation Center, Florida State Univ., Rept. 6, 1971.
18. R. L. Smith. Time-Stability Analysis of the Trip Generation and Pre-Distribution Mode-Choice Models. Univ. of Michigan, PhD dissertation, 1973.

A SYNTHESIZED TRIP-FORECASTING MODEL FOR SMALL- AND MEDIUM-SIZED URBAN AREAS

M. R. Zaryouni, Civil Engineering Division, Rensselaer Polytechnic Institute; and
E. J. Kannel, Department of Civil Engineering,
University of Illinois at Urbana-Champaign

Because of the high monetary and time costs associated with home-interview surveys for urban transportation studies, planning analysts have sought to model travel demand by using other data sources such as 1970 census work-trip data. The purpose of the research reported in this paper is to examine trip distribution functions that may be appropriate for estimating zone trip interchange in small- to medium-sized urban areas. Several functional forms of travel impedance were investigated. For the city sizes studied, model accuracy is shown to be relatively insensitive to the form of the travel impedance function. Analytical deductions are used to develop a calibration technique for a 2-way, constrained gravity model using the simple negative exponential function. Calibration of the model can be accomplished without using extensive origin-destination survey data. The model is tested by using data from actual studies, and an outline is suggested for calibrating the distribution model by using the 1970 census data.

ORIGIN-DESTINATION (O-D) data collection is the most time consuming and costly part of any transportation study. Furthermore, the level of accuracy of O-D surveys is frequently so low that the interzone trip forecasts based on them are unreliable. This occurs because the number of dwelling units in a traffic zone is small and the number of traffic zones within the system is large; therefore, the trip exchange between 2 given zones is a rare attribute of the zone population, and a very high sampling rate is required to provide acceptably accurate O-D data. Long (6) has shown that, when a city with a population of 100,000 is divided into 200 traffic zones, the error of non-home-based interzone trip exchanges based on a 5 percent home-interview sampling rate could be as high as ±270 percent. The problem is compounded further if financial constraints force a sampling rate as low as 2 percent for small- or medium-sized urban areas (2).

The cost, time consumption, and inaccuracy of O-D surveys demonstrate the merit of exploring synthesized models that can be calibrated by using available information and, therefore, do not call for conducting a particular O-D survey. Given that O-D data have inherent inaccuracies and that given the inaccuracy inherent in O-D data and that synthesizing models are structured on the state of the art, it is not clear that the predictions based on an appropriate synthesizing model would be less reliable than those based on costly and time-consuming travel surveys. In fact, the ultimate goal of transportation science may be perceived as the reaching of a mature stage where models and not particular surveys are capable of providing sufficiently accurate predictions for decision making.

This paper presents a gravity-distribution model for small- and medium-sized urbanized areas that can be calibrated by using trip-end information. The suggested calibration method eliminates the need of having extensive home-interview O-D data for the distribution stage of the trip forecasting procedure. Furthermore, the paper outlines how use of this distribution model eliminates the need for conducting an O-D study for trip forecasting in small- and medium-sized cities. The elimination of O-D surveys from transportation studies is based on 3 premises.

1. Estimates of trip productions and trip attractions of home-based work person

Selection of Distribution Function

The principal problem in developing a distribution model is the selection of the form of $F(C_{ij})$ and the quantification of the parameters of this function. The appropriate functional form of the distribution function has been the subject of many research efforts (1, 7, 13, 14). Both simple and complex functions have been suggested. However, examination of the nature of the small- and medium-sized cities reveals that using a complex distribution function for such areas is not necessary.

Because of the limited destination opportunities available to travelers in a small- or medium-sized urban area, travelers usually do not face a real choice among equivalent but locationally different opportunities. Therefore, the cost of reaching an opportunity cannot be a major factor in selecting a given opportunity by a class of travelers. Furthermore, because of the insignificance of the travel cost in such areas, travelers may not differentiate meaningfully among the cost of reaching different opportunities. A study by Zaryouni (15) shows that the consideration of travel cost provided a model only 7 to 10 percent more predictive than a gravity model with no consideration of travel cost [$F(C_{ij}) = 1$] for Billings, Montana (population 60,000), and Decatur, Illinois (population 110,000). It may be concluded that the consideration of travel cost has a marginal effect on the predictivity of a gravity model compared to trip-end information for small- and medium-sized urban areas; therefore, the predictivity of a gravity model cannot be very sensitive with respect to the functional form of the distribution function. This logical conclusion has been supported empirically by the Zaryouni study (15). The study (15) demonstrated that the inverse power and the negative exponential distribution functions provide practically the same goodness of fit for the gravity model as the more complex gamma function does. In this study, the distribution model results were compared with actual trip tables developed from traditional transportation surveys. The principal measure used as the criterion for the predictivity of a gravity model was the relative deviation d defined here as

$$d = \sum_i \sum_j (T_{ij} - S_{ij})^2/S_{ij} \tag{5}$$

where S_{ij} = interzonal trip volume from actual survey. The lower the d is, the better the goodness of fit and the more predictive the model will be. The parameter or parameters of each distribution function are determined by using an iterative procedure to minimize d.

The minimum deviation d obtained for 3 different distribution functions is given in Table 1. The table shows that the minimum deviation is practically the same for the 3 functions tested. Therefore, the negative exponential function that is a more appropriate function for calibration purposes is suggested for use in small- and medium-sized urban areas. Equation 1, then, becomes

$$T_{ij} = r_i s_j \exp(-\beta C_{ij}) \tag{6}$$

Model Calibration

The calibration problem is to estimate parameter β. One way to find β is to equate average travel cost predicted by the model $\overline{C}_m$ to actual travel cost measured from actual O-D data (4, 11). However, to do so, one would need actual average travel cost, which, in turn, would require a more extensive O-D survey. Instead, here, the value of β will be estimated by deriving a relationship between β and $\overline{C}_m$ that can be solved iteratively for β.

The average travel cost predicted by a 2-way, constrained gravity model can be written as:

$$\overline{C}_m = \frac{\sum_i \sum_j T_{ij} C_{ij}}{\sum_i \sum_j T_{ij}} = \frac{\sum_i \sum_j r_i s_j C_{ij} \exp(-\beta C_{ij})}{\sum_i \sum_j r_i s_j \exp(-\beta C_{ij})} \tag{7}$$

The product $r_i s_j$ is a function of the average travel cost of all trips that originated at i and the average travel cost of all trips that ended at j. Being a function of an overall average, the $r_i s_j$ dependency on a particular value of C_{ij} is not significant when the number of zones is relatively large (5). Therefore, $r_i s_j$ can be replaced by its expected value and can be taken outside the summation sign and cancelled from the numerator and the denominator. Equation 7 may then be written and reduced to

$$\overline{C}_m = \frac{\int_0^\infty C \exp(-\beta C) dC}{\int_0^\infty \exp(-\beta C) dc} = \frac{1}{\beta} \tag{8}$$

Equation 8 suggests a procedure for estimation of β. The value of β should be selected so that $\overline{C}_m$ equals $1/\beta$. This value of β can be obtained by using an iterative procedure. For each selected β, a trip table is computed by a 2-way, constrained gravity model. Then, the $\overline{C}_m$ associated with the selected β is computed from

$$\overline{C}_m = \frac{1}{T} \sum_i \sum_j C_{ij} T_{ij} \tag{9}$$

By using this procedure for a range of values of β, one can plot $\overline{C}_m$ against β. The intersection of the $\overline{C}_m$ curve and $1/\beta$ curve gives the optimum value of β. Figures 1 and 2 show the calibration method for the cases of Billings, Montana, and Decatur, Illinois. In these figures, the d associated with each value of β also is shown. The figures demonstrate that the value of β determined by the suggested method provides practically the minimum d for the gravity model.

A SYNTHESIZED MODEL FOR TRIP FORECASTING

This section, by using the result of the previous section, will demonstrate the possibility of trip forecasting without conducting an extensive O-D study. The aim is to suggest not a detailed procedure but a general outline. Published census tract data of 1970 provide many useful data other than those that will be mentioned. Indeed, other studies are being conducted to analyze the potential of census data for urban transportation planning (12). These research efforts are more detailed and often require significantly more manipulation and adjustment of the data. When this experience in estimating trip-generation rates from census studies is available this information should be exploited and, where appropriate, the more detailed procedures should be used.

The proposed synthesized model may be thought of in 4 parts:

1. Production of and attractions for home-based work (HBW) trips,
2. Interzone trip exchanges for HBW trips,
3. Assignment of HBW trips, and
4. Computation of peak-hour volume (PHV) and average daily traffic (ADT) for each major link.

Table 1. Minimum deviation and corresponding parameters for selected distribution functions.

Distribution Function	Billings, Montana			Decatur, Illinois		
	d	α	β	d	α	β
$F(C_{ij}) = C_{ij}^{-\alpha}$	37,026	-1.10		34,489	-1.0	
$F(C_{ij}) = \exp(-\beta C_{ij})$	36,504		0.11	34,305		0.09
$F(C_{ij}) = C_{ij}^{\alpha} \exp(-\beta C_{ij})$	36,474	-0.2	0.09	34,233	-0.4	0.06

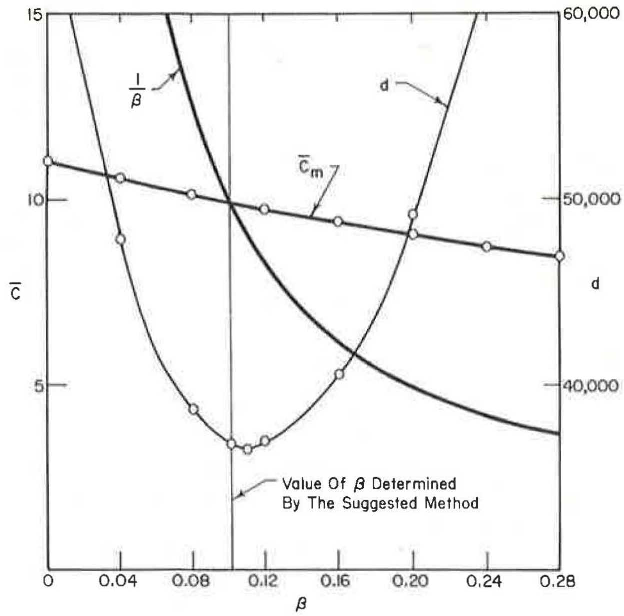


Figure 1. Model calibration relationships for Billings, Montana.

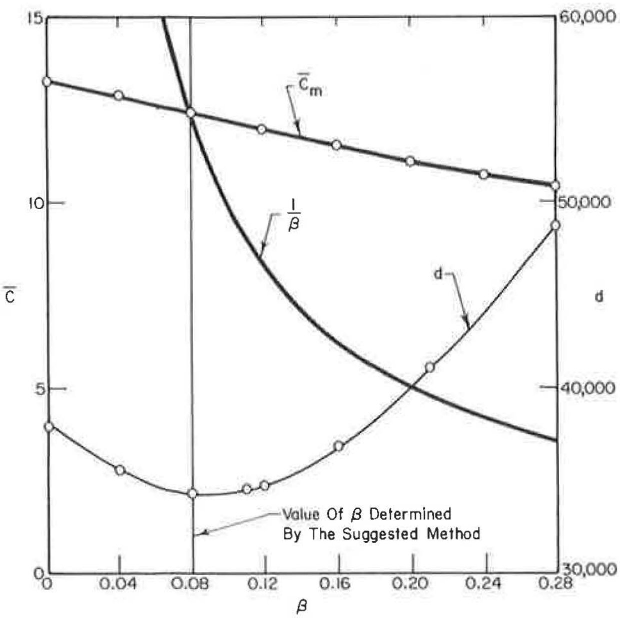


Figure 2. Model calibration relationships for Decatur, Illinois.

Production of and Attractions for Home-Based Work Trips

An HBW trip is defined as 1 trip/employee/day that originates from the place of residence and is destined for the place of work of each employee. Clearly, the total number of HBW trips for the whole study area is the same as the total number of employed people who live and work within the area.

Published census tract data provide the number of employed people who live in each census tract. This number is assumed to be the same as the number of trip productions for the HBW trips. Furthermore, census tract data of 1970 distinguish between those employees who work outside and those who work inside the standard metropolitan statistical area (SMSA). This distinction makes it possible to obtain the internal trip production for the HBW trips more accurately if the SMSA is selected as the study area. Also, in the census, the number of people employed in their residences is given for 15 industries. Therefore, stratification of the trip productions is possible.

If 1 HBW trip/employee is assumed, then the number of trip attractions for the work zones would be the number of employees in the work zone. Although the 1970 census recorded employee work address, the level of detail and accuracy is questionable. Employment location data may be supplemented, however, from other data sources such as state employment security records and the major employers. These government and private data sources provide an opportunity to stratify the attraction component of the HBW trips and to obtain information regarding work trips produced by employees outside the SMSA.

Because the production and attraction components of the HBW trips usually are not provided by the same source, the data components may not be consistent with each other. Some modification may be needed to make these components consistent. Rather than use the zone productions and attractions directly from these sources, one might be better advised to estimate the total number of HBW trips in the entire area first and then to compute the distribution of trip ends from the available data sources. Hence T equals the total number of people who live and work within the area and could be obtained now from existing data sources and later from lane use and economic forecasting. Then, if $(ER)_i$ is the total number of employed people who live at zone i and $(EW)_j$ is the total number of employed people who work at zone j, let

$$u_i = \frac{(ER)_i}{\sum_i (ER)_i} \tag{10}$$

$$v_j = \frac{(EW)_j}{\sum_j (EW)_j} \tag{11}$$

Then the production and attraction components of HBW trips become, respectively,

$$P_i = Tu_i \tag{12}$$

$$A_j = Tv_j \tag{13}$$

If the trip-end data base has been stratified according to industry, this procedure may be carried out for each industry separately.

Interzone Trip Exchanges for Home-Based Work Trips

Based on knowledge of the trip productions and attractions from the previous step, a 2-way, constrained gravity model with the negative exponential distribution function provides an interzone-trip-exchange model for the HBW trips:

$$F(C_{ij}) = \exp(-\beta C_{ij})$$

β is computed according to the previous section of this paper.

When the data have been stratified according to industries, a separate distribution choice for each industry should be followed. The results then must be added into a single interzone trip matrix for the next step.

Assignment of Home-Based Work Trips

Based on knowledge of interzone HBW trips from the previous step, one can determine the routes that travelers use when going from one zone to another and assign HBW trips to the street network on the basis of route choice. This step could be done by any existing assignment model, including the simple method of judgment.

Computing Peak-Hour Volumes and Average Daily Traffic for Each Major Link

Other researchers have examined the work trip to evaluate peak-hour and daily travel patterns (7, 10). Shunk, Grecco, and Anderson (10) have shown that a strong linear relationship exists between the HBW trips passing through a major street and the PHV and ADT of that major street. For link ℓ it may be written

$$(PHV)_\ell = K_1 \times (HBW)_\ell \tag{14}$$

$$(ADT)_\ell = K_2 \times (HBW)_\ell \tag{15}$$

or more accurately,

$$(PHV)_\ell = a + b(HBW)_\ell \tag{16}$$

$$(ADT)_\ell = a' + b'(HBW)_\ell \tag{17}$$

Some actual vehicle counts in major streets should be made (existing traffic maps may provide ADT information as well), and the PHV and ADT should be computed. Next, with the corresponding HBW from the previous step, K_1 and K_2 can be computed:

$$K_1 = \frac{1}{N} \sum_{\ell}^{N} \left(\frac{PHV}{HBW}\right)_\ell \tag{18}$$

$$K_2 = \frac{1}{N} \sum_{\ell}^{N} \left(\frac{ADT}{HBW}\right)_{\ell} \tag{19}$$

If the second more complex model is used, a, b, a′, and b′ can be derived by using an appropriate regression analysis. It should be noted that for each type of major street (freeway, arterials, and collectors) a different set of indexes preferably should be developed. Also the values of K_1 and K_2 or a, b, a′, and b′ derived for the present must be assumed to prevail in the future unless data forecasting a change are available.

SUMMARY AND CONCLUSION

The principal objective of this research was to evaluate trip-distribution models that can be calibrated with minimal travel-survey data and can be used to estimate travel patterns in small- and medium-sized urban areas. The 2-way, constrained gravity model was found to be relatively insensitive to the functional form of the distribution function for the cities studied. Therefore there is no need for oversophistication in the form of the distribution function. The negative exponential function is simple but not meaningfully less predictive than other complex functions. The calibration method for the model is based on analytical deductions. The primary advantage is that the interzone travel patterns can be estimated with limited travel-survey data.

The model is proposed to be used where trip-generation and trip-attraction estimates can be obtained primarily from census data. Although the validity of the distribution model has been tested, the total synthetic modeling approach still must be examined in an initial study.

REFERENCES

1. N. Ashford and D. Covault. The Mathematical Form of Travel Time Factors. Highway Research Record 283, 1969, pp. 30-47.
2. Decatur Urbanized Area Transportation Study, A Summary Report. Illinois Division of Highways, 1970.
3. A. W. Evans. Some Properties of Trip Distribution Methods. Transportation Research, Vol. 4, No. 1, 1969, pp. 19-36.
4. A. W. Evans. The Calibration of Trip Distribution Models With Exponential or Similar Cost Function. Transportation Research, Vol. 5, 1970, pp. 15-38.
5. H. R. Kirby. Normalization Factors of the Gravity Model—An Interpretation. Transportation Research, Vol. 4, No. 1, 1970, pp. 37-50.
6. G. D. Long. Accuracy of Dwelling Unit Survey in Travel Forecasting. Traffic Engineering, Vol. 44, No. 5, 1974, pp. 54-62.
7. P. S. Parsonson. Development of an Urban Peak-Hour Traffic Model Based on the 1970 Census and Concurrent Ground Counts. Georgia Institute of Technology, 1972.
8. L. March. Urban System: A Generalized Distribution Function. In Urban and Regional Planning, Pion Limited, London, 1971, pp. 157-170.
9. T. Sasaki. Probability Method to Estimate Trip Distribution. Annual Reports of Roads, Japan Road Association, 1967, pp. 121-134.
10. G. A. Shunk, W. Grecco, and V. L. Anderson. The Journey to Work as a Singular Basis for Travel Pattern Survey. Purdue Univ., Joint Highway Project 26, 1967.
11. J. A. Tomlin and S. G. Tomlin. Traffic Distribution and Entropy. Nature, Vol. 220, No. 5171, 1968, pp. 974-976.
12. Census Data and Urban Transportation Planning. TRB Special Rept. 145, 1974.
13. Alan M. Voorhees and Associates. Factors and Trends in Trip Lengths. NCHRP

Rept. 48, 1968.
14. A. G. Wilson. A Statistical Theory of Spatial Distribution. Transportation Research, Vol. 1, 1967, pp. 253-269.
15. M. R. Zaryouni. A General Formulation of Trip Distribution. Univ. of Illinois at Urbana-Champaign, PhD dissertation, June 1974.

DEVELOPMENT OF A SIMULATION MODEL FOR REGIONAL RECREATIONAL TRAVEL

W. D. Berg and P. A. Koushki, University of Wisconsin—Madison;
C. L. Krueger, Chicago Department of Public Works; and
W. L. Bittner, City of Antigo, Wisconsin

A simulation model for recreational travel was developed for use in analyzing the impact of outdoor recreational travel by residents of a 9-state Upper Midwest region to Michigan, Minnesota, and Wisconsin. Travel data were collected for 6,441 randomly selected households by using a telephone home-interview survey procedure. After the trips were stratified into summer-vacation and summer-weekend categories, cross-classification analysis was used to relate household trip-making frequencies to the socioeconomic characteristics of the household and its accessibility to recreational attractions within the study area. Gravity models then were calibrated to distribute recreational trips within a county-level system of demand and supply zones. Zone trip productions were estimated by applying the cross-classification model to those households within each demand zone; zone trip attractions were established synthetically on the basis of the reported trip ends from the telephone home-interview survey and the distribution of seasonal homes across the region. The gravity-model trip tables then were assigned to a regional highway corridor network for comparison with automatic-traffic-recorder data.

•TOURISM and outdoor recreation are key elements in the economy of the sparsely developed Upper Great Lakes region, which consists of the upper half of the states of Michigan, Minnesota, and Wisconsin. To help meet the growing leisure-time demand of the heavily populated Upper Midwest states, the water-based recreational potential of the area is being furthered by the development and expansion of private, state, and U.S. parks and recreational facilities. Of concern to planners, public officials, and environmentalists is the resulting increased population pressure that will be exerted on environmentally sensitive locations.

If critical land and water use problems are to be anticipated and overcome, officials at the federal, state, and local levels must have information on future levels of recreational demand. This becomes a complex, long-range planning problem because of the many factors that influence recreational travel and participation. Among these are transportation network improvement plans; state and federal recreational development plans; energy conservation policies such as reduced highway speed limits and fuel allocation programs; and the growth and spatial distribution of the regional population.

To address these problems, a multidisciplinary, recreational planning study was undertaken for the Upper Great Lakes Regional Commission (8). The research reported in this paper concerns the simulation model for regional recreational travel that was developed in the transportation phase of the study.

DATA COLLECTION

Several early constraints were placed on the development of the recreational travel simulation model because of earlier work that had been completed on the phase of the project concerning the recreational demand survey and forecasts. Most of the demand for outdoor recreation to the Upper Great Lakes region was assumed to originate in the 9-state Upper Midwest region (UMR) consisting of North Dakota, South Dakota, Minnesota, Iowa, Wisconsin, Illinois, Michigan, Indiana, and Ohio. Travel originating beyond these states would not be measured in the study. In addition, the University of Wisconsin Survey Research Laboratory had been hired to develop an interview procedure by which residents of the UMR could be questioned about their participation in outdoor recreational activities. A telephone home-interview survey was being pretested when authorization was received to proceed with the transportation phase of the project. Several final modifications to the questionnaire then were made so that it might provide certain origin-destination (O-D) information considered essential for the development of the travel simulation model. One of these was a request by the 3 state transportation agencies to consider travel to any point within the entire states of Michigan, Minnesota, and Wisconsin, the Great Lakes region (GLR).

Recreational Travel Patterns

The survey questionnaire then was applied to a sample of 6,441 households from the UMR during the late summer and fall of 1972. Households were selected from a computer-generated frame of all possible telephone numbers in the 9 states. Each respondent was to provide information about 2 types of outdoor recreational trips: those that involved at least 5 days away from home and those that involved at least 2, but not more than 4, days away from home. This essentially defined vacation trips and weekend trips respectively. Although all respondents were to submit information about each outdoor recreational trip taken by a member of the household during the 12 months preceding the interview, detailed data pertaining to their travel and activity participation were collected only for trips that occurred during the five "summer" months of May through September and that were destined to or passed through the GLR.

It was found subsequently that trips made by children under 18 years of age by themselves could not be included in the model because of a lack of sufficiently complete O-D data. Although the survey revealed a large number of person trips by children, most of these likely involved some type of group excursion such as travel to a summer camp or a weekend scouting camp-out. Trips of this nature generally would be made by bus and therefore would play a minor role in the simulation of vehicle travel.

Because each reported trip was a round trip, the return portion of each trip was assumed to follow the same path as the outbound portion and the majority of travelers were assumed to follow a somewhat direct, or minimum-path, route to and from their main destination. These assumptions were not considered overly restrictive because the objective of the study was to simulate travel over major corridors of a 3-state region, and any minimum-path routing over this network generally would encompass a relatively wide band of potential, intermediate recreational stops.

Transportation Network

Before establishing the corridor network, we defined a system of county-level traffic analysis zones. Within the 3-state GLR, each of the 243 counties represented 1 traffic analysis zone. For the remaining 6 states, 41 multicounty zones were defined on the basis of population distribution and distance from the GLR. Zone centroids were selected by considering the location of population centers, major highways, and prominent recreational facilities.

The Interstate Highway System and the major state arterial highways within the 3-state GLR were used to define the recreational travel corridors for the region. This network was more generalized than a complete statewide highway traffic assignment network would be, and yet it provided a more realistic identification of major corridor travel flows than a spiderweb network linking adjacent county centroids would provide. Although the study was concerned only with corridors in the GLR, the network was extended in a similar fashion throughout the remaining 6 states of the UMR study area. Corridor links in these states were defined for the most part by the Interstate Highway System.

Link lengths were determined from state highway maps for the various states. The average travel speed for each link was based on the location and functional classification of the corridor. If a particular link represented 2 or more parallel highways or if it was characterized by 2 or more distinct subsections having different average travel speeds, the multiple sections were weighted by their length or travel speed or both to yield a single average link. Each zone centroid then was connected to the network by 1 or more dummy links depending on the nature of the local transportation system.

Recreational Supply

Because the study was concerned with travel to the 3-state GLR, an inventory was made of the availability of recreational supply data on a county-level basis for the region. Although detailed data were maintained by each state, there was a lack of uniformity in the type of information that was recorded and the units of measurement that were used. As an alternative, the Public Outdoor Recreation Areas and Facilities Inventory (9) undertaken in 1972 was used. Although this survey could provide detailed data on state- and U.S.-administered recreational facilities in each county, no comparable data set was available for the extensive supply of privately operated recreational facilities throughout the GLR.

As a result, the problem of measuring the recreational attractiveness of a county was approached on a generalized basis by assuming that privately owned and privately operated recreational facilities were likely to predominate in those counties that also possessed certain natural recreational resources. If the total attractiveness of a county could be estimated by simply measuring natural recreation resources, such as lakes and major public recreational areas, then detailed recreational supply data would be unnecessary. Therefore, the total area of state- and U.S.-administered parks plus the total area of lakes was established as the county-level recreational supply variable. Water area for those counties bordering one of the Great Lakes included an area of that lake equivalent to the length of its shoreline times 0.5 mile (0.8 km).

TRIP GENERATION

The recreational trip making frequencies of the sample dwelling units are given in Table 1. It is important to note that, for all categories of vacation and weekend recreational trips, most respondents made either 1 trip or no trips. Looking specifically at summer trips to the GLR, one finds that only 1.5 percent of the households made more than 1 vacation trip and that only 4.1 percent made more than 1 weekend trip. This had the effect of making recreational trip production a dichotomous variable.

Although subsequent phases of the recreational travel simulation model would require zone trip productions as input data, the small O-D survey sampling rate (approximately 0.05 percent) precluded the development of a zone trip production model. Instead, effort was directed toward the formulation and testing of a household trip production model based on the 6,441 dwelling units that had been surveyed. The frequency of summer recreational trips per dwelling unit to the GLR was assumed to be a function of the socioeconomic characteristics of the household, the relative supply of outdoor recreational facilities available to the household, and the relative travel cost associated with taking a trip to a recreational attraction in the GLR.

The recreational supply and travel cost parameters for the trip-generation model were formulated as an accessibility index:

$$AI_i = \sum_{j=1}^{m} \left[\frac{S_j}{t_{ij}^{1.5}}\right] \tag{1}$$

where

AI_i = accessibility of a dwelling unit in production zone i to the recreational supply of the GLR,
m = total number of attraction zones in the GLR,
S_j = total area of lakes plus state and U.S. parks in attraction zone j, and
t_{ij} = minimum-path travel time from the centroid of zone i to the centroid of zone j.

Because the small sample size used in the telephone home-interview survey did not permit the building of a base-year O-D trip table, computing a set of gravity-model friction factors for the travel time function could not be done. The value of 1.5 for the travel-time exponent was selected on the basis of a correlation analysis between the trip-rate variables and alternative formulations of the accessibility index.

Cross-classification analysis then was selected as the technique by which summer recreational trip productions to the GLR would be related to household socioeconomic characteristics and recreational-attraction accessibilities. Because the accessibility index represented a zone characteristic, all dwelling units in a given zone were assumed to have the same accessibility to recreational attractions in the GLR. To compare and evaluate the several models that were to be tested, an analysis of variance was performed on the trip-rate variable for each model. Consideration also was given to the ease with which the distributional characteristics of the independent variables could be measured and forecast on a county-level basis.

Of the several cross-classification trip-production models that were formulated, the 2 selected for use in the travel-simulation model expressed summer-vacation trips and summer-weekend trips as a function of family income, occupation of head of household, and accessibility to the recreational attractions of the GLR. The complete sets of tables for both models are available from the authors. Each model has the same structure: 5 classes of family income, 10 classes of occupation of head of household, and 5 classes of accessibility index.

Figure 1 shows the relationship between family income and recreational trip rate. As income increases to $25,000/year, both vacation and weekend outdoor recreational trip rates increase. This supports the frequently observed tendency for higher income families to make more trips of all types. The decline in both trip rates for families with annual incomes of more than $25,000 probably is due to their ability to take a more expensive type of trip than one involving outdoor recreation in the GLR.

Figure 2 shows the relationships between occupation of head of household and recreational trip rate. The figure shows that those occupations that traditionally provide higher incomes are associated with higher trip rates. However, the fact that occupation also reflects the nature of the leisure time available to the head of the household possibly is of greater importance. People in occupations associated with high trip rates (especially those who are professionals or are in business) ordinarily can set their own working hours. This would not be the case for people who are employees. Although the farmer ordinarily is self-employed, his or her free time is confined generally to the winter months, which results in the fact that summer trip rates for farm families are the lowest for all occupation categories.

Figure 3 shows the relationship between accessibility to recreational attractions and trip rate. As accessibility to the recreational areas of the GLR increases, the likelihood of a family's making a recreational trip to the area also increases, particularly for weekend trips in which a high premium is placed on minimizing travel time. This relationship underlines the importance of an efficient transportation system to those regions whose economies are tied closely to tourism and outdoor recreation.

The coefficients of determination (r^2's) for the vacation-trip and weekend-trip models are 0.11 and 0.17 respectively. Although zone-level trip-production models have yielded r^2 values above 0.90, they have been shown to be equivalent to household-level trip-production models with r^2 values near 0.25 because of the nature of the variance that is being explained (10). As a demonstration of the effect of level of aggregation on the variance explanation of recreational trip rates, a set of zone-level trip-production models were formulated. Multiple linear regression analysis then was used to relate

Table 1. Percentage of sample dwelling units that took summer recreational trips in 1972.

Type of Trip	Number of Trips 0	1	2	3	4	5	6
Vacation							
All trips	64.3	28.4	6.0	1.0	0.2	0.1	–
Trips to GLR	87.2	11.2	1.2	0.3	–	–	–
Weekend							
All trips	72.8	18.9	5.7	1.7	0.9	0.3	0.3
Trips to GLR	85.3	10.7	2.5	0.8	0.5	0.2	0.1

Figure 1. Family income versus trip rate.

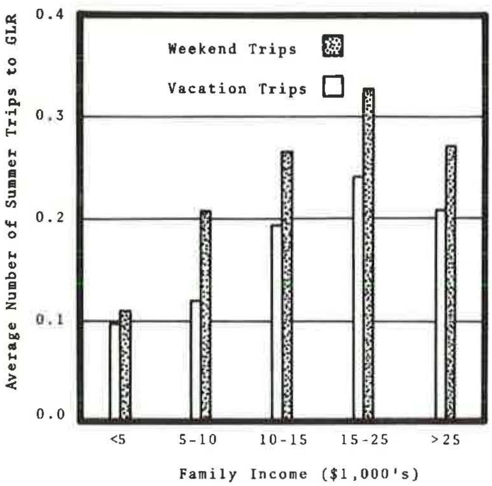


Figure 2. Occupation of head of household versus trip rate.

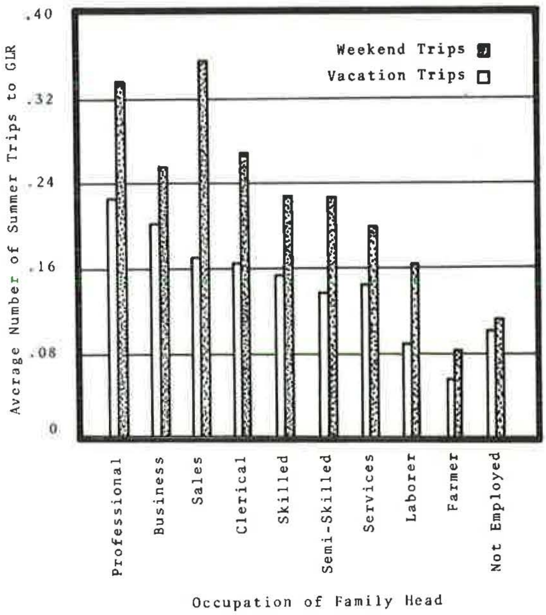
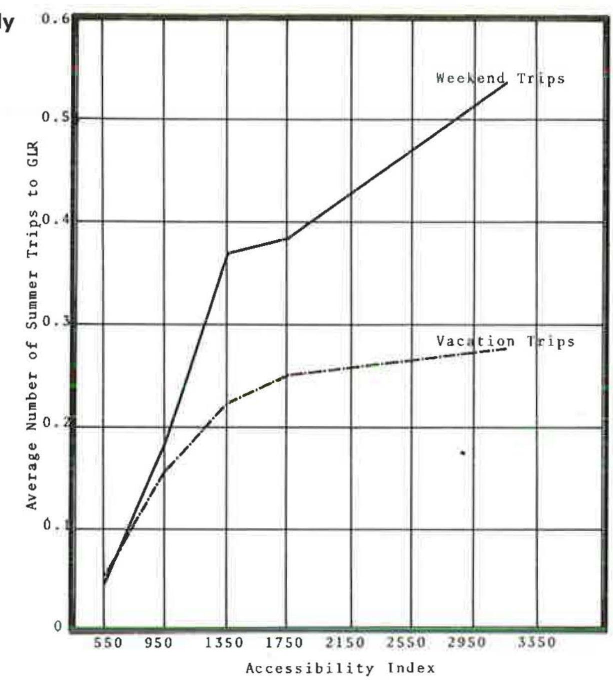


Figure 3. Accessibility to recreational supply versus trip rate.



the 1970 estimated zone trip productions to the total population of zone i (POP_i), average family income of zone i (INC_i), and AI_i of the zones. Models for vacation and weekend trips were tested in log form, and a summary of them is given in Table 2. The regression statistics reveal that, when the within-zone trip-production variance is removed, 78 to 92 percent of the remaining between-zone variance can be explained by the selected zone characteristics.

The basically unstable nature of outdoor recreational trips is another explanation of the low variance of the cross-classification models. Work trips are predictable on an hourly and daily basis, but a household's recreational travel can exhibit a large variation in destination from year to year even though it may begin on a certain day of the week or month with some consistency. This characteristic creates a built-in variance within a recreational travel model that has been estimated to be as much as 20 percent (2).

TRIP DISTRIBUTION

The gravity model was selected for allocating zone recreational trip productions to the 243 county-level recreational supply zones in the GLR.

Trip Productions and Attractions

Because only 0.05 percent of the dwelling units in the UMR had been interviewed (the number of summer recreation trips to the GLR totaled 961 vacation trips and 1,386 weekend trips), many zones were found to have no reported trip ends whatsoever. As a result, base-year trip productions and attractions had to be developed synthetically. Summer-vacation and weekend trip productions for each of the 284 UMR zones were estimated by applying the cross-classification model to 1970 household data. The resulting 1970 zone trip productions for the 9-state UMR totaled 1.59 million vacation trips and 2.34 million weekend trips over the 5-month period of May through September. Base-year trip attractions for the 243 zones in the GLR then were established by using a 2-stage procedure. First, total base-year trip productions were allocated to each of 36 districts proportional to the trips reported by the survey respondents:

$$A_k = \left(\sum_i P_i\right) \left[\frac{D_k}{\sum_k D_k}\right] \tag{2}$$

where

A_k = number of summer-vacation or weekend recreational trips attracted to district k,
P_i = number of summer-vacation or weekend recreational trips produced in zone i, and
D_k = mean number of reported summer-vacation or weekend recreational trips to district k.

Second, the trip-attraction total of each district was allocated to the individual zones comprising each district on the basis of the number of seasonal homes in the zones within a given district:

$$A_j = A_k \left[\frac{S_{jk}}{\sum_j S_{jk}}\right] \tag{3}$$

where

A_j = estimated zone trip attractions, and
S_{jk} = number of seasonal homes in zone j within district k.

Although the availability of seasonal home data was not discovered until late in the study, it proved to be a conceptually sound and statistically significant indicator of the recreational attractiveness of an area.

The relationship between A_j and selected zone characteristics then was tested in log form by using multiple linear regression analysis. Independent variables included number of seasonal homes SH_j, total area of lakes plus state and U.S. parks REC_j, and accessibility of the zone to the population of the UMR AI_j. The accessibility variable was computed as

$$AI_j = \sum_{i=1}^{n} \left[\frac{POP_i}{t_{ij}^{1.5}}\right] \tag{4}$$

where n = total number of zones in UMR. The resulting regression models for vacation and weekend zone trip attractions are tabulated as follows (level of significance is 5 percent):

Type of Trip	Model	Number of Observations	r^2
Vacation	$\log A_j = 0.78 \log SH_j + 0.35 \log REC_j$	243	0.99
Weekend	$\log A_j = 0.78 \log SH_j + 0.42 \log AI_j$	243	0.99

The regression statistics indicate that the models explain 99 percent of the variance in the estimated trip attractions for the 243 zones in the GLR. In comparing the 2 models, one finds that the number of seasonal homes in a zone has a significant influence on both vacation and weekend recreational trip attractions. Also vacation trips tend to be attracted to those zones with extensive natural recreation resources, and weekend trips tend to be attracted to those zones that are most accessible to regional population concentrations. These relationships demonstrate a trade-off between the drawing power of major recreation areas and the impedance of travel time. For extended recreational trips, people are willing to spend more time traveling to reach prominent outdoor recreational attractions. For weekend recreational trips, however, travel time becomes more important than the character of the recreational supply at the point of destination.

Gravity-Model Calibration

After the estimated number of base-year (1970) zone trip productions and zone trip attractions for both vacation and weekend recreational trips were established, 2 sets of gravity-model friction factors were calibrated by using 30-min travel-time intervals (5). The calibrated friction factors for both vacation and weekend recreational trips are shown in Figure 4. A comparison of the associated trip-length frequency distributions shown in Figures 5 and 6 again reveals the previously noted influence of travel time on recreational-travel patterns. The mean trip length for summer-vacation trips is 261 min. For summer weekend trips, it is 170 min. This is a difference of 1.5 hours or approximately 90 miles (144 km) of additional 1-way travel.

Because of the structure of the gravity-model trip-interchange equation, large devia-

Table 2. Zone trip-production models for summer recreational trips to the Great Lakes region.

Type of Trip	Population	Model	Number of Observations	r^2
Vacation	<500,000	$\log P_i = -6.61 + 1.52 \log INC_i + 0.82 \log POP_i$	259	0.78
	>500,000	$\log P_i = -10.5 + 1.37 \log INC_i + 1.02 \log POP_i + 1.08 \log AI_i$	24	0.92
Weekend	<500,000	$\log P_i = -8.11 + 1.12 \log INC_i + 0.84 \log POP_i + 1.02 \log AI_i$	259	0.90
	>500,000	$\log P_i = -12.4 + 1.72 \log INC_i + 0.96 \log POP_i + 1.39 \log AI_i$	24	0.89

Note: Level of significance is 10 percent.

Figure 4. Calibrated gravity-model friction factors.

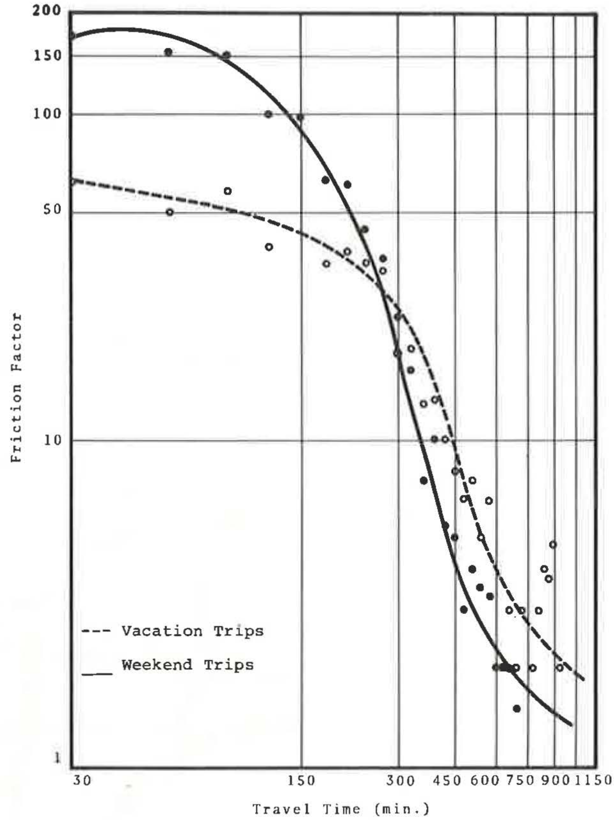
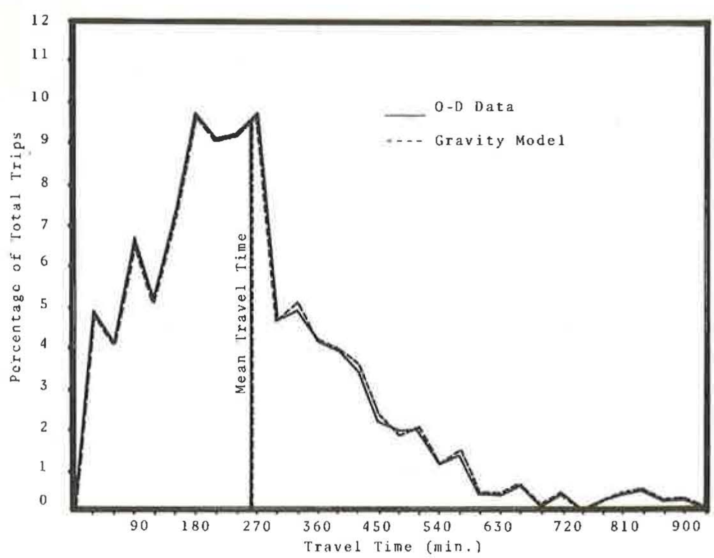


Figure 5. Gravity model versus origin-destination trip-length distributions for vacation trips.



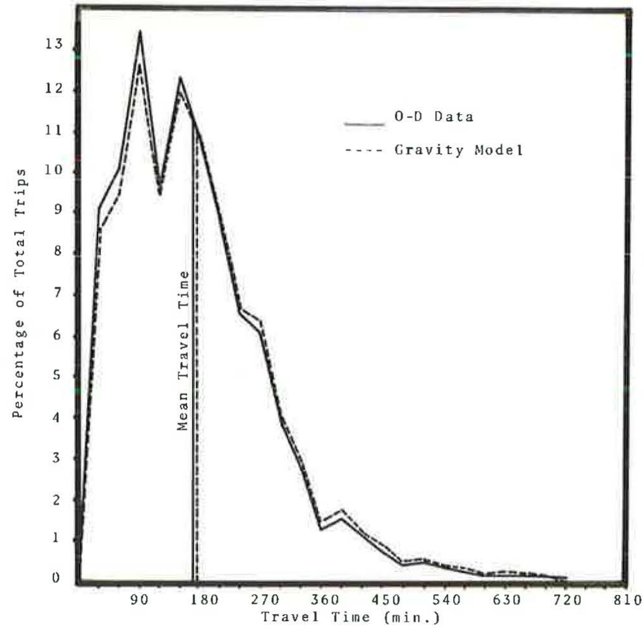


Figure 6. Gravity model versus origin-destination trip-length distributions for weekend trips.

Figure 7. Simulated 1970 peak-weekend-day recreational traffic flow for Wisconsin.

tions existed between the synthesized trip attractions and the distributed trips when they were compared on a district level. The zone trip attractions therefore were adjusted and another iteration of the gravity model was performed (10). The results of the second iteration revealed that a reasonably good balance had been achieved, although a slight increase in trip length did occur. Because those districts that received too few trip ends generally were in the more remote parts of the region and those that received too many trip ends were located close to major metropolitan centers such as Chicago and Detroit, it was judged that further iterations of the model would yield only an improved trip-end balance at the expense of an imbalance in trip-length distribution.

Therefore, the simulated trip interchanges from the second iteration were accepted as a reasonable and sufficient estimate of base-year recreational travel to the GLR.

TRAFFIC ASSIGNMENT

In the final stage of the simulation model, the combined trip table for the estimated 1970 summer recreational trips to the GLR was assigned to the corridor network by using an all-or-nothing assignment (10). The trip table represented 1-way summer recreational trips from zone of residence to zone of main destination. On the basis of the peaking and modal characteristics of the trips reported in the telephone home-interview survey, the assigned trips were factored to represent the peak nondirectional vehicle-travel flow on an average weekend day in July or August. Figure 7 shows the resulting estimated 1970 recreational traffic flow pattern for Wisconsin.

An important aspect of the entire recreational travel simulation process was the testing of the link assignments. The usual statistical accuracy criteria could not be applied directly because of a lack of traffic count data for recreational travel. As an alternative, a series of comparisons were made between the assigned link volumes and the total traffic counts at cut lines defined by selected automatic-traffic-recorder stations in each of the 3 states.

A total of 87 stations located on the Interstate Highway System and the state highway systems were identified as being compatible with a particular link of the corridor assignment network. The average Saturday and Sunday volumes for July and August then were averaged for each station and compared with the corresponding link volumes for an average weekend day in July and August. The results of the cut-line analysis revealed that the assigned link volumes generally varied from 5 to 60 percent of the total traffic count at the automatic-traffic-recorder stations. Those links representing routes with a high functional classification tended to have a higher percentage of recreational traffic than did other corridors offering a lower level of mobility, which supports the assumption that recreationists want to minimize the travel costs incurred in reaching their final destination.

Some of the assigned link volumes in the extreme northern part of the region were relatively low because travel to Canada was not incorporated in the model. Inspection of traffic-flow maps for the region revealed sizable volumes of traffic on routes that lead directly into Canadian provinces. Subsequent modeling and assignment of these trips would increase the percentage of recreational traffic along these corridors to the expected higher levels.

APPLICATION OF THE SIMULATION MODEL

Although the cut-line analysis was quite subjective, the recreational travel simulation model was considered to offer a reasonable level of accuracy for system-level corridor planning purposes in the GLR. Furthermore, because the cross-classification trip-production models and the multiple regression zone trip-attraction models are sensitive to the level of service provided by the regional transportation system, analyses can be made of the impact of alternative transportation investment and pricing policies on levels of recreational travel demand. When analyzing specific highways within a corridor, one can make the corridor flow proportional to parallel highways through the use of travel-time diversion curves. However, these flows will reflect only overnight recreational travel from place of residence to point of main destination and return.

An additional use of the corridor flow data is in the analysis of the direction and magnitude of external traffic entering and leaving a given local study area. This can be of aid in the planning and development of major thoroughfares that provide access to prominent recreational attractions. However, the trip-attraction models were developed by using synthesized trip-end estimates for the established zones. This approach was followed because the telephone home-interview survey procedure could not provide a sufficient sample of trips for direct expansion to total trip attractions.

Although measures of privately operated recreational attractions, quality of recreational facilities, or degree of crowding within the county-level zones could not be included, trip-attraction forecasts nevertheless can be adjusted subjectively to account for such factors.

CONCLUSION

The experience gained during the course of this study indicates that the application of traditional urban transportation planning methods can offer a workable approach to the analysis of statewide recreational travel. However, additional research is needed in a number of areas. For example, the most efficient methods of collecting recreational travel data need to be established. This problem is related closely to the need to achieve a better understanding of recreational travel behavior, especially trip frequency and choice of destination. Little is known about the role played by promotional campaigns, degree of crowding, and quality of recreation in attracting travelers to various locations. And how recreational and nonrecreational travel models can be integrated to provide a composite view of total statewide travel patterns needs to be established.

ACKNOWLEDGMENTS

This research is based on work completed for the Upper Great Lakes Regional Commission as a part of an Upper Great Lakes Regional Recreation Planning Study, Technical Assistance Project, which was funded through the University of Wisconsin Institute for Environmental Studies and directed by the University of Wisconsin-Extension.

The cooperation and suggestions of the Michigan Department of State Highways and Transportation, the Minnesota Department of Highways, and the Wisconsin Department of Transportation during the course of the study are gratefully acknowledged. Special appreciation also is extended to the Division of Planning, Wisconsin Department of Transportation, for computer-programming consultation and services.

REFERENCES

1. J. A. Deacon, J. G. Pigman, K. D. Kaltenbach, and R. C. Deen. Models of Outdoor Recreational Travel. Highway Research Record 472, 1973, pp. 45-62.
2. J. B. Ellis. A Systems Model for Recreational Travel in Ontario: A Progress Report. Ontario Department of Highways, Rept. RR 126, 1967.
3. J. S. Matthias and W. L. Grecco. Simplified Procedure for Estimating Recreational Travel to Multi-Purpose Reservoirs. Highway Research Record 250, 1968, pp. 54-69.
4. J. G. Pigman, R. C. Deen, and J. A. Deacon. Characteristics of Outdoor Recreational Travel. Highway Research Record 472, 1973, pp. 28-44.
5. D. C. Robinson and W. L. Grecco. Stability of Recreational Demand Model. Highway Research Record 392, 1972, pp. 147-156.
6. R. L. Smith and E. D. Landman. Recreational Traffic to Federal Reservoirs in Kansas. Engineering Experiment Station, Kansas State Univ., Aug. 1965.
7. E. C. Sullivan, R. D. Layton, and A. Kanafani. Transportation Analysis Techniques for National Forest Planning. Highway Research Record 408, 1972, pp. 17-25.
8. Upper Great Lakes Regional Planning Study, Part 3: Transportation. Institute for Environmental Studies, Univ. of Wisconsin-Extension, Madison, 1974.
9. Public Outdoor Recreation Areas and Facilities Inventory. Bureau of Outdoor Recreation, U.S. Department of the Interior.
10. Urban Transportation Planning—General Information. Federal Highway Administration, U.S. Department of Transportation, March 1972.

DISAGGREGATE MULTIMODAL MODEL FOR WORK TRIPS IN THE NETHERLANDS

Moshe Ben-Akiva,
Massachusetts Institute of Technology and Cambridge Systematics, Inc.; and
Martin G. Richards, Buro Goudappel en Coffeng, b.v., Netherlands

This paper describes a disaggregate modal-choice model with 6 travel modes (walking, bicycle, moped, car, bus, and train) for work trips. The data used in the study were from 2 communities adjacent to Eindhoven, Netherlands. A number of alternative model specifications were tested, and the results of these tests were analyzed. The model specification that was considered to be the most satisfactory overall is based on treating in-vehicle travel time as a generic variable and out-of-vehicle travel time as a series of modal-specific variables. Out-of-pocket travel costs were found to have no significant influence on modal choice. Although a number of socioeconomic variables were tried, the only ones included in the most satisfactory model were 3 vehicle-availability variables (car, moped, and bicycle). Analyses of the coefficients estimated in a number of different subsamples of the main sample showed that the marginal decreases in the standard errors of the coefficients were very small for samples containing more than 250 observations. A simple test of the most satisfactory model estimated as an aggregate predictive model indicated that effects of the theoretical problems of using a disaggregate model with aggregate data can be minimized by use of suitable market stratification.

•THIS PAPER presents some of the results of a study to develop disaggregate, behavioral travel demand models in the Netherlands. The general disaggregate modeling methodology and the multinomial logit model used in this study are described in detail elsewhere (1, 3, 4) and will not be repeated here. The purpose of this paper is to describe the models that were developed for work trips.

The models predict the short-term modal-choice behavior of a worker; residential location, work place, and automobile ownership are considered as predetermined. The models, therefore, predict the choice of travel mode for a given work trip. Other choices that received some attention in this study were between (a) a single trip to work and a double work chain that is taken by workers who go home for lunch and (b) the choice of both destination and mode for shopping trips (3, 4, 13).

The urban transportation scene in the Netherlands is characterized by the number of modes available and in common use; bus, car, bicycle, moped, and walking all play a significant part. In the larger urban areas, the train also is used for intraurban trips. In Amsterdam, The Hague, Rotterdam, and Delft, there is an extensive streetcar network; the streetcar network in Rotterdam is supplemented by a metro. In 1966, 9.4 percent of all weekday trips in the west of the country were made by public transport. Buses and streetcars accounted for 7.6 percent, and trains accounted for 1.8 percent (7). Although 10.7 percent of all trips made within the 4 major urban areas were by bus or streetcar, only 1.6 percent of all weekday trips made in the other urban areas were by public transport. In these other urban areas, bicycle and moped predominated and accounted for 55.6 percent of all weekday trips. Even within the 4 major urban areas, bicycle and moped were used for 35.7 percent of all trips.

The bicycle is a well-established and well-known Dutch phenomenon. The moped is more recent, and the number in use has doubled between 1960 and 1971. The moped is a motorized bicycle with an engine of not more than 50-cm^3 capacity; it can be ridden by anyone 16 years old or older. Officially, mopeds are limited to 30 km/h within built-up areas and 40 km/h elsewhere. In practice, these speed limits are not enforced. Under congested urban conditions, because mopeds have a high degree of maneuverability, journey times of mopeds are considerably shorter than those of bus and equal to those of car. Passengers are allowed to be carried, and the Dutch Central Bureau of Statistics estimated that, in 1971, mopeds were used for 9.77 million passenger-km; private cars were used for 78.4 million passenger-km (5). Thus a conventional binary modal-split model for car and public transport clearly has only limited usefulness, and the bicycle, moped, and walking modes also must be modeled explicitly.

THEORETICAL MODEL

The choice-of-mode-to-work model explains the conditional probabilities of choosing a mode of travel for the work trip given residential and employment locations and given that a trip is made. Thus the dependent variable can be denoted as follows:

$$P(m:M_t) \tag{1}$$

or

$$P_t(m) \tag{2}$$

where

m = an alternative mode, and
M_t = set of available modes for traveler t.

The logit model predicting this probability is written as

$$P(m:M_t) = \frac{e^{V_{mt}}}{\sum\limits_{m' \in M_t} e^{V_{m't}}} \tag{3}$$

where V_{mt} is the utility of mode m to traveler t for the work trip and can be expressed in the general form

$$V_{mt} = V_m(Z_m, S_t) \tag{4}$$

where

Z_m = a vector of characteristics of mode m, and
S_t = a vector of socioeconomic characteristics of traveler t.

V_{mt} is assumed to be a linear function in the parameters.

$$V_{mt} = X'_{mt}\,\theta = \sum_{k=1}^{K} X_{mtk}\,\theta_k \quad (5)$$

where

X_{mt} = a $K \times 1$ vector of finite functions constructed from the various Z_m and S_t variables and different from one alternative to another = $(X_{mt1}, X_{mt2}, \ldots, X_{mtK})$, and
θ = a $K \times 1$ vector of coefficients to be estimated for each model = $(\theta_1, \theta_2, \ldots, \theta_K)$.

If a variable appears only in the utility function of mode m, then it is a mode-m-specific variable that takes a value of 0 in all other modal utilities. If a variable appears in the utility function of all modes, then it is a generic variable. The value of a generic variable must not be equal for each alternative (that is, each mode) for all observations, or, mathematically, this variable is canceled out.

DATA

The data for the estimation of the models were taken from a 1970 home-interview travel survey conducted in Eindhoven, the fifth largest city in the Netherlands (1970 population: 190,000), and 4 adjacent municipalities: Best, Veldhoven, Geldrop, and Son and Breugel (2, 12). One particularly relevant characteristic of these data was that the origin-destination (O-D) data had been coded to a 10-m rectangular coordinate system. This made identifying the precise locations of the O-D addresses possible. For this study only, the data for residents of Best and Son and Bruegel were used. Son and Breugel (1970 population: 10,800) is a medium- and high-income area with bus but no train service to Eindhoven; Best (1970 population: 16,500) is a low-income area with both rail and bus links to Eindhoven.

The trip and socioeconomic data available from the survey were supplemented by transportation level-of-service data collected for this study. Level-of-service data were derived by manually locating each pair of home and work addresses on large-scale plans and by using a variety of information sources including original measurements.

The sample used for estimation was limited to home-work-home chains as opposed to simple home-work trips. This was done to minimize the influence of other choice considerations such as use of car for business purposes during the day. The sample included 390 observations of a single home-work-home chain during the day.

For model evaluation and practical considerations, the sample was divided into 2 different sets of subsamples. One was a random division into 2 equal groups based on whether the reference number of the household was an odd or an even number; these 2 groups were coded SBB1 and SBB2. The other was a geographic division into Son and Breugel (SB) and Best (B). The total set of data was coded SBB. This division was intended to allow evaluation of model stability, by comparing the 2 random subsamples, and geographic stability.

The household and personal socioeconomic data were examined to determine whether, on the basis of household vehicle ownership and possession of a driver's license, car, moped, or bicycle could be regarded as alternative modes. If no vehicles were owned in any of the 3 classes (car, moped, bicycle) then that mode was assumed not to be a valid alternative. Similarly, if the individual was not in possession of a valid driver's license for cars, even if a car was available within the household, then car was not considered to be a valid alternative. (Only car driver was considered as an alternative; car passengers were excluded.) If a destination was more than 2 km away, walk was not considered a relevant alternative mode. Train was considered relevant only from Best to Eindhoven. Bus was not considered a relevant alternative for work trips within Son and Breugel because no such trips by bus were observed.

Table 1 gives the distribution of the chosen mode for work trips in the SBB sample. This is compared with the modal split of weekday trips for all purposes in the west of the Netherlands (7).

VARIABLES USED IN THE MODELS

The explanatory variables used in the work modal-choice models are of 2 types: level-of-service and socioeconomic variables. We have denoted the variables by abbreviations. Table 2 gives a list of the variable codes and their descriptions. The following are the code prefixes and what they represent:

Prefix	Meaning	Prefix	Meaning
B	Bus	PT	Public transit
BF	Moped	T	Train
C	Car	TW	Two-wheel vehicle
F	Bicycle	W	Walking

Level-of-Service Variables

Three level-of-service variables were used in the models.

1. IVTT represents in-vehicle travel time (in minutes). For walking trips, IVTT is always 0; for all mechanical modes it is the time spent in or on the vehicle.
2. OVTT represents out-of-vehicle travel time (in minutes). For walking trips, OVTT is the total walking time of the trip. For car, bicycle, and moped, it is denoted as POVTT, which is defined as the time taken to walk to and from the parked vehicle, bicycle, or moped as well as to park and unpark. Bus and train OVTT consists of 2 parts: WSOVTT and SOVTT. WSOVTT is defined as the time spent walking to and from the bus stop or station. SOVTT is the time spent at a bus stop or station as well as in transferring from a bus to a train or vice versa.
3. OPTC represents out-of-pocket travel cost (in Dutch cents). For walking and bicycle trips, OPTC has a value of 0. For car and moped trips, it has a value equal to fuel costs in keeping with the traditional expectations of perceived motoring costs; no parking charges were included. For bus and train, OPTC equals the costs of the fares.

Socioeconomic Variables

Seven socioeconomic variables were used in the models.

1. HHINC represents annual household income. Annual income data were coded according to the following 6 classes (in guilders): (a) less than 5,000, (b) 5,001 to 25,000, (c) 10,001 to 15,000, (d) 15,001 to 20,000, (e) 20,001 to 25,000, and (f) more than 25,000.
2. PER represents number of persons in the household 5 years old or older.
3. AOD represents number of private cars and noncommercial vans reported as owned by the household divided by number of licensed drivers in the household. AOD was not permitted to have a value of more than 1.0.
4. BOP represents number of bicycles reported as owned by the household divided by number of persons 5 years old or older in the household.
5. MOA represents number of mopeds reported as owned by a household divided by number of persons 15 years old or older in the household.
6. HHPOS represents position in household. This variable equals 1 for head of household and 0 for others. Because the purpose of this variable in the modal-choice

model was to represent car availability, it also was assigned a value of 1 for adults with drivers' licenses who were not heads of households if there was perfect car availability, that is, if AOD for that household was equal to 1.0.

7. OCC represents occupation of traveler. This was used as a simple dummy variable taking a value of 1 for professionals, managers, and executives and 0 for others.

A number of other variables available from the original data files, such as age and sex, were considered as was a more detailed description of the variable OCC, but these were excluded during the course of the work because of deductive considerations or simply because of the limited number of different specifications that could be estimated.

Socioeconomic variables do not vary across alternatives. Because of the form of the model, these variables somehow must be transformed either by combining them with other variables or by making them alternative-specific (that means including them in the utility function of some modes and not in others or allowing their coefficients to vary across modes).

Modal Constants

Modal constants have a totally different function than the other variables have. If the variables included in the modal utility functions fully explain modal-choice behavior, then the modal constants, or more generally, the pure alternative effects, should equal 0. Thus, with a perfect model specification and with perfect data, it can be argued that no constants are necessary. However, estimating a model without constants is not recommended in practice because the estimated values of the coefficients of the variables included could be seriously affected if those variables do not explain fully the observed behavior. The constants therefore represent the effect of those variables that influence modal choice but are not included explicitly in the model. The formulation of the logit model is such that constants have to be alternative-specific, or, in this case, modal-specific.

If we have reason to believe that those variables that should have been included in the model to make it complete were excluded and have different values for different situations, then the values of the constants also will differ. Under such circumstances, the use of a model estimated on data for one area to predict behavior in another area, at a different time, or for a different socioeconomic group may be questionable. In a modal-choice model, the modal constants partially represent travelers' evaluations of level-of-service variables, such as reliability, comfort, privacy, and convenience, which are either difficult or impossible to measure, and unobserved preferences of travelers. An attempt was made in this study to account for the pure alternative effects through the introduction of various modal-specific socioeconomic and vehicle-availability variables that usually could be expected to be highly correlated with the unobserved variables. The exclusion of constants then was considered acceptable if the coefficients of the various level-of-service variables were not affected significantly.

In a model with a maximum of 6 alternatives, only 5 (alternative-specific) constants, or coefficients of a given socioeconomic variable, can be identified. The walking mode therefore was used as the base alternative and the coefficients of modal-specific variables, such as income, should be interpreted relative to the walking mode. The constants and the coefficients of socioeconomic variables introduced for the public transit modes were combined for bus and train primarily because of the small number of trips by both of these modes.

ALTERNATIVE SPECIFICATIONS

The initial estimation runs were deliberately restricted to rather simple specifications and were initially done with only half of the total sample (SBB1) because level-of-service data for the full SBB sample were not available until later in the study.

The results of these initial runs indicated that OPTC (whether or not it was divided

Table 1. Chosen modes.

Mode	Work Trips in Son and Breugel and Best Sample		Weekday Trips in West of the Netherlands (percent)
	Percent	Observations	
Car	40	156	23
Bicycle	28	108	30
Moped	20	79	11
Bus (and streetcar)	5	21	8
Train	3	11	2
Walk	4	15	22
Other	—	—	4

Table 2. Summary of variable codes.

Code	Description	Alternative to Which Applicable
BFCON	Constant	Moped
BFHHINC	Household income	Moped
BFHHINC/PER	Household income divided by number of persons	Moped
BFMOA	Moped ownership (number of mopeds divided by number of persons 15 years old or older)	Moped
BFOCC	Occupation	Moped
BFOVTT	Total out-of-vehicle travel time	Moped
CAOD	Car availability (number of cars divided by number of licensed drivers)	Car
CHHINC	Household income	Car
COCC	Occupation	Car
COVTT	Total out-of-vehicle travel time	Car
FBOP	Bicycle ownership (number of bicycles divided by number of persons 5 years old or older)	Bicycle
FCON	Constant	Bicycle
FHHINC	Household income	Bicycle
FHHINC/PER	Household income divided by number of persons	Bicycle
FOCC	Occupation	Bicycle
FOVTT	Total out-of-vehicle travel time	Bicycle
HHINC	Household income	
HHINC/PER	Household income divided by number of persons in household	
IVTT	In-vehicle travel time	All except walking
OPTC	Out-of-pocket travel costs	Car, moped, bus, and train
OVTT	Total out-of-vehicle travel time	
POVTT	Parking out-of-vehicle travel time	Car, bicycle, and moped
PTCON	Constant	Bus and train
PTHHINC	Household income	Bus and train
PTHHINC/PER	Household income divided by number of persons	Bus and train
PTOCC	Occupation	Bus and train
SOVTT	Waiting and transfer time	Bus and train
TWOCC	Occupation	Bicycle and moped
WOVTT	Walking time	Walking
WSOVTT	Walking time to and from bus stop or station	Bus and train

Table 3. Coefficients and standard errors of in-vehicle travel time estimated for SBB 1 sample.

Mode	Coefficient	Standard Error	Mode	Coefficient	Standard Error
Car	-0.0997	0.0584	Bus	-0.759	0.0297
Bicycle	-0.0995	0.0258	Train	-0.0881	0.0580
Moped	-0.1273	0.0516			

by income) had a small positive coefficient, which was not significantly different from 0. This means that OPTC does not significantly influence modal choice, at least in this case. This result corresponds with a deductive assumption made in a modal-choice study with data from Amsterdam and Rotterdam (8) in which OPTC was assumed to have no influence on the choice of whether to drive a car to work in existing Dutch urban transportation conditions. The results obtained from this study could, however, be explained by the small differences between the costs of the different modes because of the generally low level of transport costs and the relatively short length of the trips in the sample.

Previous modal-choice models have tended to introduce level-of-service characteristics as generic variables (1, 6, 10). In this study, this practice was not necessarily justified for the various components of OVTT. This is one of the few studies ever conducted with totally disaggregate data (data in which the home and work addresses could be located precisely), and thus the walking time to and from bus stops or stations could be appraised accurately. Parking time, however, could only be estimated because no information was available on the location of the parking place; waiting and transfer times also could only be estimated. Thus the different components of OVTT were estimated with varying degrees of reliability; aggregation to give total OVTT therefore was avoided in most of the specifications tried. The work by Stopher, Spear, and Sucher (14) on the measurement of inconvenience in urban travel suggests, however, that a division of OVTT into its various components would be preferable to aggregation into a single variable.

An exploratory run on the SBB1 sample in which IVTT was used as a modal-specific variable indicated that IVTT reasonably could be applied as a generic variable because little difference was found between the modal-specific coefficients as shown by the data given in Table 3.

From the results of some of the early models estimated, HHPOS and OCC were observed to have no significant influence on modal choice for the sample of data used for this study.

In general, the variables were introduced into the modal utility functions in a linear form. Only in a few cases was an attempt made to explore nonlinear finite functions of the variables; this was done with the time variables and AOD, which was only used as a car-specific variable CAOD. For the time variables, a natural log transformation was tried. For the CAOD variable, 2 finite functions were tried; one was a product of CAOD and HHINC, and the other was the product of CAOD and the natural log of IVTT by car.

The estimation results for the alternative specifications tried with the full SBB sample (390 trips) are given in Table 4; models 5 to 14 are based on the same specification for the level-of-service variables. The differences among models 5 to 14 are in the specifications of the various alternative-specific constants and the socioeconomic and vehicle availability variables. Table 5 gives for the model data in Table 4 the ℓn likelihood function of $\theta = 0$ (which corresponds to all alternatives being equally likely), $L^*(0)$, and the ℓn likelihood function of $\theta = \hat{\theta}$, $L^*(\hat{\theta})$, where $\hat{\theta}$ is the vector of estimated coefficients. Table 5 also gives the statistic $-2\,[L^*(0) - L^*(\hat{\theta})]$, which is asymptotically distributed as χ^2 with degrees of freedom equal to the number of coefficients. ρ^2 is a measure of goodness of fit; it is equal to $1 - \frac{L^*(0)}{L^*(\hat{\theta})}$, or the ratio of the explained ℓn likelihood to the total ℓn likelihood, and it lies between 0 and 1. $\bar{\rho}^2$ is ρ^2 adjusted for degrees of freedom.

DISCUSSION OF ESTIMATED COEFFICIENTS

Values of Estimated Coefficients

The strongest deductive knowledge that we have about the estimated values of the coefficients is on their signs. We expect that, with everything else held equal, a deterioration in the level of service offered by any mode will reduce the probability of that mode's being chosen. Thus an essential requirement is that the utility of any one mode should decrease as the value of most level-of-service variables increases. (This is not the case, of course, with a level-of-service variable such as comfort if comfort was measured on a scale that increased with increasing comfort.) If a given level-of-service variable enters a utility function only once, then, with the exception of some specific transformations, the coefficient of that variable can be expected to be negative. If, however, the variable is entered in more than 1 form, such as a simple variable and in a logarithmic transformation, then it is possible that only 1 of the coefficients need be negative. Thus in model 3 (Table 4), for instance, the sum of a (-0.14 OVTT + 0.52 ℓnOVTT) decreases with increasing values of OVTT, if OVTT is greater than 3.5 min (OVTT was always greater than 3.5 in the data set used for estimation); in fact, in this particular case, the coefficient of ℓnOVTT was not significantly different from 0. This general requirement is satisfied in all the models estimated except for OPTC in every model in which it was included. Because the coefficient of OPTC was never found to be significantly different from 0, OPTC was assumed not to have any significant influence on modal choice in this particular sample; therefore, OPTC was ultimately excluded.

There are also some deductive expectations with respect to the relationships between

Table 4. Estimation results of alternative specifications for SBB sample.

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
Variable or Constant	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
IVTT	-0.0673	0.0091	-0.0757	0.0106	-0.0672	0.0116			-0.0644	0.0101	-0.0732	0.0101	-0.0664	0.0107
ℓn IVTT					-0.0343	0.5016	-1.5639	0.3971						
OVTT	-0.1193	0.0172			-0.1419	0.0346	-0.1610	0.0234						
ℓn OVTT					0.5186	0.7366								
WOVTT			-0.1676	0.0291					-0.1661	0.0321	-0.1545	0.0293	-0.2535	0.0535
POVTT									-0.2933	0.0794	-0.2372	0.0727	-0.2314	0.1018
COVTT			-0.3290	0.0803										
FOVTT			-0.2501	0.1015										
BFOVTT			-0.5141	0.1040										
WSOVTT			-0.1206	0.0210					-0.1195	0.0218	-0.1149	0.0209	-0.1131	0.0222
SOVTT			-0.0674	0.0236					-0.0962	0.0344	-0.0887	0.0329	-0.0836	0.0341
OPTC/HHINC	0.0066	0.0101			0.0086	0.0115	0.0384	0.0100	0.0081	0.0107	0.0119	0.0105	0.0067	0.0113
CHHINC													-0.5690	0.2616
CAOD	0.5877	0.4396	1.1363	0.0528									0.7647	0.5048
CAOD × HHINC					0.2408	0.1170	0.4968	0.1162	0.2435	0.1418				
CAOD × HHINC/PER											0.2963	0.4183		
CAOD × ℓn IVTT														
COCC														
FCON	0.3254	0.3514			0.7486	0.6009	1.4463	0.5795						
FHHINC									0.0286	0.1138			-0.5294	0.2422
FHHINC/PER											0.2480	0.3527		
FBOP														
FOCC													0.3037	0.5665
BFCON	-0.7845	0.3517			-0.3864	0.4712	0.3563	0.3673						
BFHHINC									-0.2562	0.1081			-0.8109	0.2411
BFHHINC/PER											-1.1043	0.3626		
BFMOA														
BFOCC													-0.1390	0.6262
TWOCC														
PTCON	1.8615	0.5976			1.8984	0.8316	3.1672	0.7681						
PTHHINC									0.1099	0.1823			-0.4557	0.2707
PTHHINC/PER											0.1670	0.4585		
PTOCC													0.1654	0.6759

	Model 8[a]		Model 9		Model 10		Model 11		Model 12		Model 13		Model 14	
Variable or Constant	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
IVTT	-0.0649	0.0105	-0.0643	0.0102	-0.0644	0.0103	-0.0665	0.0094	-0.0715	0.0090	-0.0721	0.0092	-0.0600	0.0093
ℓn IVTT														
OVTT														
ℓn OVTT														
WOVTT	-0.2527	0.0535	-0.2511	0.0530	-0.2201	0.0527	-0.1257	0.0266	-0.1333	0.0266	-0.1095	0.0281	-0.1192	0.0295
POVTT	-0.2255	0.1013	-0.2145	0.0907	-0.2539	0.0920	-0.3559	0.0786	-0.3590	0.0795	-0.2675	0.0919	-0.3260	0.0946
COVTT														
FOVTT														
BFOVTT														
WSOVTT	-0.1113	0.0220	-0.1101	0.0213	-0.1041	0.0214	-0.0987	0.0193	-0.1034	0.0193	-0.1269	0.0238	-0.1136	0.0234
SOVTT	-0.0834	0.0341	-0.0828	0.0340	-0.0936	0.0345	-0.0853	0.0306	-0.0511	0.0216	-0.0824	0.0289	-0.0856	0.0288
OPTC/HHINC	0.0071	0.0113	0.0071	0.0113	0.0210	0.0119	0.0189	0.0113						
CHHINC	-0.5764	0.2616	-0.5899	0.2559	-0.5866	0.2567								
CAOD	0.7721	0.5034	0.7636	0.5024	1.5314	0.5444	2.1651	0.4753	2.2928	0.4709	2.2025	0.4690		
CAOD × HHINC														
CAOD × HHINC/PER														
CAOD × ℓn IVTT													1.0056	0.1800
COCC	-0.1213	0.4928												
FCON														
FHHINC	-0.5332	0.2421	-0.5393	0.2408	-0.6860	0.2448								
FHHINC/PER														
FBOP					1.5790	0.3841	1.4698	0.3399	1.3272	0.3263	1.4936	0.3410	1.4348	0.3211
FOCC														
BFCON														
BFHHINC	0.8209	0.2409	-0.8236	0.2406	-0.8091	0.2530								
BFHHINC/PER														
BFMOA					1.8492	0.7457	2.2659	0.5558	-0.0132	0.5375	0.2281	0.5527	0.6689	0.5536
BFOCC	0.0197	0.4537												
TWOCC	0.0497	0.4537												
PTCON											1.7404	0.9155	1.5057	0.9252
PTHHINC	-0.4594	0.2696	-0.4553	0.2690										
PTHHINC/PER														
PTOCC	0.1086	0.5897												

[a]3 different versions of model 8 were estimated, each of which incorporated only 1 of the 3 variables COCC, TWOCC, and PTOCC; all of the coefficients, except for those of TWOCC and PTOCC, and standard errors are those for the model in which COCC was included.

Table 5. Likelihood functions and other data for models in Table 4.

Model	L*(0)	$L^*(\hat{\theta})$	χ^2	ρ^2	$\bar{\rho}^2$
1	-435.13	-272.81	324.63	0.37	0.37
2	-435.13	-271.69	326.88	0.38	0.37
3	-435.13	-271.17	327.92	0.38	0.37
4	-435.13	-292.60	285.05	0.33	0.32
5	-435.13	-269.14	331.97	0.38	0.37
6	-435.13	-272.08	326.10	0.37	0.37
7	-435.13	-266.55	337.16	0.39	0.38
8	-435.13	-266.83	336.59	0.39	0.38
9	-435.13	-266.86	366.53	0.39	0.38
10	-435.13	-257.59	355.07	0.41	0.40
11	-435.13	-268.69	332.88	0.38	0.38
12	-435.13	-270.10	330.04	0.38	0.37
13	-435.13	-268.29	333.68	0.38	0.38
14	-435.13	-260.54	349.17	0.40	0.40

certain coefficients of level-of-service attributes. For example, one would expect the coefficient for an OVTT variable to be greater than the coefficient for IVTT. In all the specifications tried, IVTT was, indeed, found to have a lower coefficient than any of the OVTT variables; the coefficients of walk as a mode WOVTT and WSOVTT are almost equal and have approximately twice the value of the coefficient for IVTT. Thus walking appears to be twice as inconvenient as riding in or on a vehicle. This relationship corresponds to the usual assumption used to create generalized costs in Wilson-type models in the United Kingdom, where the coefficient for excess time is usually taken to be twice that for IVTT (9). Wait time at a bus stop or station appears to be more inconvenient than IVTT but not so burdensome as walking. This is a departure from the usual assumption previously mentioned as well as from U.S. studies in which the coefficient for wait time usually is taken to be 2.5 times that of in-vehicle time (11). However, the relatively low coefficient of wait time in this sample might be attributable to a highly reliable transport service and therefore to an overestimate of the value of wait time. It also could reflect other errors in estimating the wait times used in this study. It also should be noted that, compared with the coefficients of other level-of-service variables, the coefficient for SOVTT has a relatively large variance. This probably is due to the low variability of headways in the public transit services available to the individuals in the sample (that is, low variability in SOVTT values).

Expectations with respect to the values of both the constants and the coefficients of the socioeconomic variables are more complicated and rely on very limited or nonexistent past experience. One particular problem encountered in the design of the study was the limited availability of results from previous studies using similar models especially for European circumstances.

If everything is equal, one would expect that as car ownership increases the probability of choosing car as the mode of transport would increase and thus that the probability of choosing other modes would decrease. One would therefore expect that the coefficient for CAOD would be positive and, for similar reasons, that the coefficients of both bicycle and moped availability also would be positive.

Household income and modal constants appear in the utility function of more than 1 mode, and, therefore, interpretation must relate to their relative values and not their absolute values. In model 5 (Table 4), for example, the relative values of the coefficients of household income indicate that, as household income increases and everything else is held constant, the increase in the probability of using car is relatively greater than that of using other modes, and the probability of choosing a moped will decrease relative to all other modes. Between these extremes are (a) public transit, which decreases in relation to car but increases in relation to the other modes; (b) bicycle, which decreases relative to car and public transit and increases relative to walking and moped; and (c) walking, which is a base mode. Thus the probability of walking increases relative to moped but decreases relative to all other modes; this could reflect the socioeconomic status of a moped as a transit mode.

The modal constants and socioeconomic variables could be interpreted as representing the pure preferences for the alternative modes if the utility derived from the level-of-service characteristics was equal across all modes. A direct interpretation of this kind is easier when all the level-of-service characteristics are introduced as generic variables. For example these variables in model 1 (Table 4) imply that, if IVTT and OVTT are equal across modes, then an individual with perfect car availability (AOD = 1) will rank the modes in the following order: public transit, car, bicycle, walking, and moped. The place of public transit in this pure ranking appears to be too high. However, the pure ranking for a person with all modes perfectly available, which is given by model 13 (Table 4) (which represents an improved specification), is: car, public transit, bicycle, moped, and walking. This preference ordering agrees more closely with deductive expectations; this ranking also is implied by model 14 (Table 4), which was the last specification estimated with this sample and which we considered to have the most satisfactory specification of the models estimated.

Stability and Statistical Reliability of Estimated Coefficients

The reliability and stability of the estimated coefficients can be observed in several ways including the relative magnitudes of the standard errors and the variability of the estimates across both different specifications and different subsamples.

The magnitude of the standard errors of the estimated coefficients is relatively small (compared with the magnitude of the estimated coefficients) for the travel-time variables but is relatively higher for some modal constants and socioeconomic variables, particularly the moped-specific variables (Table 4). This also could be observed from the variability of the estimated coefficients of the same variables across different specifications.

From Table 4, the coefficients of the travel-time variables are quite stable, in particular the coefficient of IVTT. On the other hand, some of the coefficients of the socioeconomic variables and the constants appear to be less stable. This pattern also was observed in the comparison of different subsamples.

As recorded in the section on data, 2 types of subsamples were created from the SBB sample. One was a random division into SBB1 and SBB2, a division that proved of value in the evaluation of the statistical stability of the estimated coefficients compared with the estimated values of the standard errors. The second was a division into 2 selected geographic subsamples, SB and B. This allowed a comparison of coefficients between 2 different areas; differences between coefficients from the 2 areas were used to trace possible specification errors. We assumed that travel behavior in both areas is similar and, therefore, that a specification that has similar coefficients for both areas is superior to a specification with divergent estimates.

Table 6 gives the estimation results for models 5, 13, and 14 for the different subsamples. From this table, one can see that the coefficients estimated for model 14 are quite stable between the SBB and B samples. The variability between the estimates of the coefficients for SBB1 and SBB2 is considerably higher, but this can probably be attributed to the fact that the standard errors are larger because of smaller sample sizes. This pattern of decreasing standard errors with increasing sample size for the 3 models and the subsamples as shown by the data given in Table 6 is shown in Figures 1, 2, and 3. These figures indicate that an increase in sample size beyond 300 observations does not reduce significantly the standard errors of the coefficients of most of the variables. This pattern suggests that, for these models, a desirable sample size would be between 300 and 400 observations.

To compare the stability of the estimated coefficients between 2 independent random samples, one would need at least 600 observations rather than the 400 observations available. Samples larger than 300 to 400 observations, however, might be necessary if more or other socioeconomic variables were to be included. With the existing sample, these have been found to have very large standard errors in contrast to coefficients of level-of-service variables for which reasonable levels of reliability at smaller sample sizes were achieved.

Geographical comparison of the estimated coefficients between B and SB is not conclusive because of the small number of observations in the SB sample. However, the stability between SBB and B seems satisfactory in view of the fact that, although B is included in SBB, there are still significant differences between both the means and distributions of several of the explanatory variables.

Table 7 gives likelihood functions and other data for the information in Table 6.

ANALYSIS OF ALTERNATIVE SPECIFICATIONS

The specification for model 14 appears to be the most satisfactory of all those tried. The reasoning leading up to this conclusion, as well as aspects of some of the other models estimated, is discussed in this section.

As discussed in the section on alternative specifications, the formulation of the level-of-service variables in models 5 to 14 seems to be superior to that used in models 1 to 4. The coefficients of all the level-of-service variables in models 5 to 14 have the ex-

Table 6. Estimation results for data sets for models 5, 13, and 14.

		SBB		SBB1		SBB2		B		SB	
Model	Variable or Constant	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
5	PTHHINC	0.1099	0.1823	0.3866	0.3205	-0.0633	0.2362	-0.0955	0.2274	-0.0528	0.6001
	FHHINC	0.0286	0.1138	0.3703	0.1777	-0.2722	0.1678	-0.2247	0.1628	0.2830	0.1980
	BFHHINC	-0.2562	0.1081	0.0265	0.1711	-0.4835	0.1576	-0.4710	0.1563	-0.0351	0.1775
	CAOD × HHINC	0.2435	0.1418	0.5316	0.2341	0.0739	0.1945	-0.0685	0.1860	0.6278	0.2766
	IVTT	-0.0644	0.0101	-0.0739	0.0155	-0.0573	0.0140	-0.0689	0.0117	-0.0501	0.0265
	WOVTT	-0.1661	0.0321	-0.1365	0.0526	-0.2028	0.0447	-0.2075	0.0426	-0.2046	0.1322
	POVTT	-0.2933	0.0794	-0.2356	0.1225	-0.3675	0.1140	-0.3418	0.0969	-0.2382	0.1597
	WSOVTT	-0.1195	0.0218	-0.0887	0.0306	-0.1591	0.0333	-0.1333	0.0272	-0.1102	0.0519
	SOVTT	-0.0962	0.0344	-0.1803	0.0650	-0.3968	0.0361	-0.0704	0.0380	-0.1047	0.1200
	OPTC/HHINC	0.0081	0.0107	0.0278	0.0175	-0.0062	0.0137	-0.0071	0.0144	0.0335	0.0268
13	IVTT	-0.0721	0.0092	-0.0836	0.0146	-0.0664	0.0123	-0.0769	0.0110		
	WOVTT	-0.1095	0.0281	-0.1074	0.0494	-0.1148	0.0349	-0.1098	0.0313		
	POVTT	-0.2675	0.0919	-0.3038	0.1459	-0.2699	0.1247	-0.2732	0.1106		
	WSOVTT	-0.1269	0.0238	-0.0849	0.0331	-0.1776	0.0378	-0.1374	0.0284		
	SOVTT	-0.0824	0.0289	-0.1121	0.0518	-0.0585	0.0339	-0.0788	0.0307		
	CAOD	2.2025	0.4690	3.4796	0.8296	1.5391	0.6052	1.4234	0.5504		
	FBOP	1.4936	0.3410	2.3249	0.5569	0.9527	0.4518	1.3537	0.3989		
	BFMOA	0.2281	0.5527	1.3823	0.8503	-0.5096	0.7479	0.1808	0.6329		
	PTCON	1.7404	0.9155	1.2698	1.5845	2.2079	1.1724	1.7068	0.9961		
14	IVTT	-0.0600	0.0093	-0.0679	0.0148	-0.0556	0.0127	-0.0688	0.0122	-0.0338	0.0208
	WOVTT	-0.1192	0.0295	-0.1210	0.0532	-0.1253	0.0366	-0.1164	0.0324	-0.2381	0.1621
	POVTT	-0.3260	0.0946	-0.3753	0.1580	-0.3464	0.1279	-0.3241	0.1130	-0.3785	0.1939
	WSOVTT	-0.1136	0.0234	-0.0701	0.0317	-0.1668	0.0378	-0.1269	0.0280	-0.0881	0.0549
	SOVTT	-0.0856	0.0288	-0.1187	0.0511	-0.0583	0.0335	-0.0828	0.0308	-0.0594	0.0856
	CAOD × ℓn IVTT	1.0056	0.1800	1.5746	0.3570	0.7718	0.2221	0.7254	0.2145	1.6427	0.4336
	FBOP	1.4348	0.3211	2.2351	0.5483	0.9502	0.4213	1.3529	0.3724	1.4551	0.6911
	BFMOA	0.6689	0.5536	1.8694	0.8818	-0.0314	0.7496	0.5000	0.6267	1.3234	1.4021
	PTCON	1.5057	0.9252	1.0927	1.5822	1.8657	1.1876	1.5553	1.0075	0.4079	2.6117

Figure 1. Standard error of estimated coefficients and sample size for model 5.

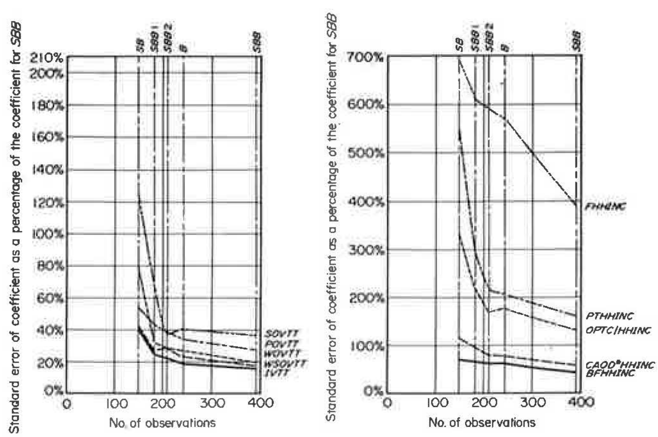


Figure 2. Standard error of estimated coefficients and sample size for model 13.

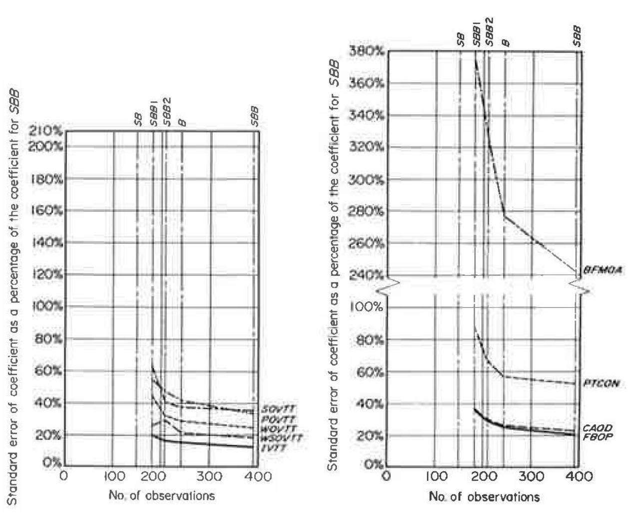


Figure 3. Standard error of estimated coefficients and sample size for model 14.

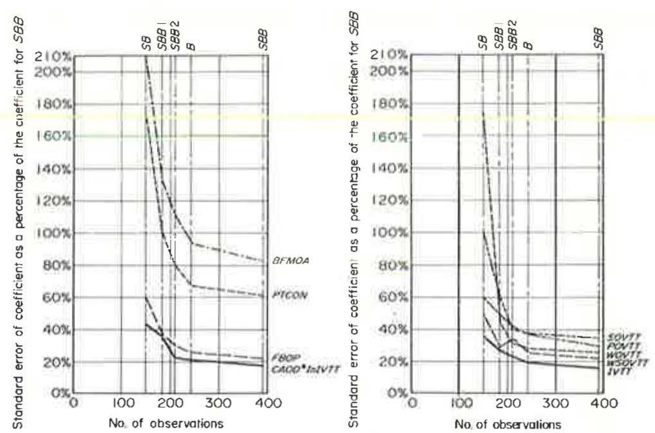


Table 7. Likelihood functions and other data for models in Table 6.

Model	Data Set	Number of Observations	$L^*(0)$	$L^*(\hat{\theta})$	χ^2	ρ^2	$\bar{\rho}^2$
5	SBB	390	-432.13	-269.14	331.97[a]	0.38	0.37
	SBB1	182	-201.98	-109.53	184.89[a]	0.46	0.44
	SBB2	208	-233.15	-149.72	166.86[a]	0.36	0.34
	B	241	-282.25	-186.92	190.65[a]	0.34	0.33
	SB	149	-152.88	-73.09	159.57[a]	0.52	0.50
13	SBB	390	-435.13	-268.29	333.68[b]	0.38	0.38
	SBB1	182	-201.98	-106.26	191.45[b]	0.47	0.46
	SBB2	208	-233.15	-153.41	159.47[b]	0.34	0.33
	B	241	-282.25	-188.96	186.58[b]	0.33	0.32
14	SBB	390	-435.13	-260.54	349.17[c]	0.40	0.40
	SBB1	182	-201.98	-101.43	201.11[c]	0.50	0.49
	SBB2	208	-233.15	-149.92	166.44[c]	0.36	0.34
	B	241	-282.25	-186.02	192.47[c]	0.34	0.33
	SB	149	-152.87	-68.29	169.18[c]	0.55	0.54

[a]Degrees of freedom = 10. [b]Degrees of freedom = 9. [c]Degrees of freedom = 9.

Table 8. Disaggregate prediction results.

Sample	Sample Size	Type of Data	Percentage of Shares: Walking	Car	Bicycle	Moped	Bus	Train
SBB	390	Predicted	2.75	40.01	27.60	21.43	5.66	2.54
		Observed	3.85	40.00	27.69	20.26	5.38	2.82
SBB1	182	Predicted	2.88	40.50	29.13	19.10	5.55	2.85
		Observed	2.75	42.86	29.67	18.13	3.30	3.30
SBB2	208	Predicted	2.65	39.57	26.27	23.48	5.76	2.27
		Observed	4.81	37.50	25.96	22.12	7.21	2.40
B	241	Predicted	3.03	31.93	31.29	22.89	6.43	4.11
		Observed	5.39	29.46	31.12	22.82	6.64	4.56
SB	149	Predicted	1.95	53.07	21.64	18.92	4.42	0
		Observed	1.34	57.05	22.15	16.11	3.36	0
Zones 100, 103, and 110 to Eindhoven center	37	Predicted	0	44.92	6.96	14.09	15.03	18.99
		Observed	0	48.65	2.70	13.51	16.22	18.92
Zones 100, 103, and 110 to Eindhoven elsewhere	74	Predicted	0	45.26	15.50	24.92	10.80	3.52
		Observed	0	37.84	17.57	28.38	10.81	5.41
Zones 200 and 210 to Eindhoven center	19	Predicted	0	61.30	12.35	14.25	12.11	0
		Observed	0	68.48	5.26	15.79	10.53	0
Zones 200 and 210 to Eindhoven elsehwere	52	Predicted	0	59.47	14.81	19.26	6.45	0
		Observed	0	63.46	19.23	13.41	3.85	0
Zone 100 total	86	Predicted	2.61	53.78	22.16	15.17	6.28	0
		Observed	2.33	59.30	22.09	11.63	4.65	0
Zone 210 total	23	Predicted	1.61	49.43	26.15	21.72	1.09	0
		Observed	0	56.52	17.39	26.09	0	0

Table 9. Aggregate prediction results.

Sample	Sample Size	Type of Data	Percentage of Shares: Car	Bicycle	Moped	Bus	Train
Best to Eindhoven, zone 2	46	Predicted	66.85	8.43	20.37	2.78	1.59
		Observed	41.30	26.09	30.43	0	2.17
Best to Eindhoven, zone 3	40	Predicted	60.35	4.48	17.83	8.98	8.45
		Observed	45.00	2.50	15.00	17.50	20.00
Son and Breugel to Eindhoven, zone 2	27	Predicted	64.03	8.95	25.22	1.78	0
		Observed	62.16	18.92	13.51	5.41	0
Son and Breugel to Eindhoven, zone 3	25	Predicted	70.96	6.72	20.16	2.16	0
		Observed	64.00	8.00	20.00	8.00	0

pected signs (except for OPTC, which is not significant) as well as the expected relative values. Therefore, any preference among specifications 5 to 14 must be based on the behavior of the socioeconomic variables and the model constants.

Models 5 and 6 have identical specifications except for household income, which was included in model 5 but replaced in model 6 by income per person. It would seem reasonable that the pure modal preferences would be more closely related to the total household income rather than the average income per person after income pooling. Although model 5 has a slightly better goodness of fit than model 6 has, the estimation results are by no means conclusive evidence that model 5 is, indeed, any better than model 6. However, this, together with the previous statement, caused us to consider model 5 superior to model 6.

In models 7, 8, and 9, the car-specific variable CAOD × HHINC was split into 2 separate variables CAOD and CHHINC. In addition, an attempt was made in models 7 and 8 to introduce OCC as a variable (in the form of a modal-specific variable), but in neither model were any of its coefficients significantly different from 0. Model 9 was considered to be less satisfactory than model 5 because of the relatively larger variance of the coefficient of CAOD; this is probably attributable to a high level of collinearity between car availability and household income.

The specification of model 10 includes the same variables as model 9 with the additional variables bicycle BFOP and moped availability BFMOA. In model 10, however, the variances of the coefficients of the socioeconomic variables were large, and thus it appeared desirable to select only a subset of these variables for the following models. This was done in models 11, 12, and 13 in which only the vehicle-availability variables were included. These variables were selected because it seemed reasonable to assume that they have a greater direct bearing on modal choice than household income has. Furthermore, the coefficients of these variables in earlier models were more significant than those of the income variables.

Models 11 and 12 are identical except for OPTC, which was excluded from model 12. In model 13, a public transit constant PTCON was reintroduced. Because it is highly probable that the specification of model 12 was not perfect, and that, therefore, the absence of a public transit constant could considerably affect the values of the coefficients of other variables, model 13 was considered the best of the series of models 7 to 13.

Thus far, therefore, models 1 to 13 have been evaluated and models 5 and 13 have been selected as 2 of the best models. These 2 models represent 2 essentially alternative specifications of the socioeconomic variables. Model 5 is based on modal-specific income variables, and model 13 is based on modal-specific vehicle-availability variables. These 2 models also were estimated for the various subsamples and the results of these estimations are given in Table 6. A comparison of these 2 sections of Table 6 and a comparison of Figures 1 and 2 show that it is evident that model 13 is more stable than model 5. From a consideration of both goodness of fit and the significance of the coefficients, one can conclude that model 13 is superior to model 5.

Examination of the variability of the estimated coefficients for the different subsamples of model 13 in Table 6, however, shows the coefficient of CAOD to be particularly unsatisfactory. Examination of the characteristics of the various subsamples showed that the coefficient of CAOD had a smaller value in those subsamples having a shorter average trip length and that it had a larger value in subsamples with a longer average trip length. It therefore seems reasonable to assume that, when a car is perfectly available to an individual, the longer the trip the more likely he or she is to choose the car. When a car is not perfectly available, and there is some degree of competition among different users of a car within a household, it would seem reasonable to expect that those individuals making longer trips will tend to have priority in use of the car over those making shorter trips. Thus it seems reasonable to specify the coefficient of car availability as a function of trip length. If this is so, then it could be expected that this function will show a diminishing marginal effect with increasing trip length and therefore the CAOD variable was multiplied by the natural log of IVTT by car. This change was implemented in model 14.

A comparison of model 13 with model 14 (Table 6 and Figures 2 and 3) shows that model 14 is superior to model 13 in terms of stability of the coefficients, significance of the coefficients, and goodness of fit. Thus, model 14 appears to have the most satisfactory specification of those models estimated and given in Table 4.

PREDICTION TESTS

Two types of prediction tests were applied. The first was a disaggregate test designed primarily to determine how well the model fit the observed data. The second test was with aggregate data and was designed to test the applicability of the model for aggregate predictions.

Disaggregate Predictions

In disaggregate predictions, explanatory variables are used to predict individual modal-choice probabilities. These individual probabilities are summed across a group of travelers and compared with the observed modal shares for the same set of individuals. When the group of travelers consists of the complete set of individuals used in the estimation of the model, this test can be viewed as a test of goodness of fit. However, the estimation procedure used in this study guarantees that, if a model specification includes a constant, the modal shares calculated in such a test will be perfect for the alternative to which that constant relates. For model 14, this implies that a disaggregate prediction test with the complete data set used in model estimation will reproduce perfectly the total public transit share because PTCON was included. However, because no constant was used for any of the remaining 4 modes, this test is still meaningful for the split between the 2 public transit modes and the 4 other modes. The results of this test are given in the first row of Table 8, and as has been stated, one can see that the split between public transit and the other 4 modes is reproduced perfectly. The results among the individual modes within these two sets, however, also are extremely satisfactory.

To provide a thorough test of the model, one would ideally apply it to a second set of data not used in the model estimation. Unfortunately, because of budget and time constraints pertaining to this study, a second set of data was not available. Instead, the test was applied to several subsets of the data set used for model estimation; the results of these also are given in Table 8. The subsets used include the subsamples SBB1, SBB2, SB, and B and 2 other types of subsamples that were created. The first of these consisted of groups of individuals with a home address in a specific zone and the second was formed of groups of individuals with a home address in a specific zone and a work address either in the center of Eindhoven or elsewhere in Eindhoven.

All the modal shares of these various subsets were predicted satisfactorily. The differences between the observed shares and the predicted shares are minimal for the larger subsets (more than 100 observations). For the smaller subsets, the relative differences between the observed and predicted shares are especially large for bicycle and moped. There is, however, a tendency for some mutual compensation here in that, when one share is overestimated, the other is usually underestimated. This means that the total share of bicycle and moped is more satisfactorily reproduced than the individual shares are.

Aggregate Predictions

In normal predictive work, disaggregate data are not usually available, and thus the model must be applied by using aggregate data. Simple substitution of group averages for the explanatory variables will result in a biased forecast of the average probability or share; this bias will disappear only if all individuals in the group for which predictions are being prepared are identical in terms of the values of all the explanatory variables. Between the 2 extremes of disaggregate predictions and use of averages for the entire group in aggregate predictions, identifying a stratification scheme or system of market segmentation so that the aggregation bias can be assumed to be small and, within the context of the application, negligible is possible.

Table 9 gives the results of aggregate predictions prepared with model 14 for 4 O-D pairs by using a very coarse zoning system. For this application of the model, all the travelers making a trip between any pair of zones were grouped together, and the average value of each variable for the members of that group was used. Furthermore, no account was taken of different sets of available modes for individuals in the group. The calculated shares are clearly not satisfactory in comparison with the observed shares. In particular, the share of car trips is significantly overestimated for the first 2 groups, which consist of people residing in Best. Although the car-share forecasts for the Son and Breugel groups are better, they are not satisfactory. The divergence between the observed and predicted shares serves as a good illustration of the errors that can occur from simply taking averages for each individual variable and directly applying them in the model.

In Son and Breugel, the majority of the residents are car owners; in Best, many families are without cars. Therefore, the model evidently performed better when there were high levels of car ownership than when there were lower levels of car ownership. Through the use of one set of average values of all variables, including CAOD, applied to all travelers, cars have effectively been made available to people for whom they were not available.

The forecasting error can be reduced by a stratification of the travelers between any pair of zones. Of course, the ultimate stratification is that of complete disaggregation, but this would also produce some prediction errors, the magnitude of which would depend on the validity of the model itself. The errors in the predictions given in Table 9 are therefore a combination of both disaggregate prediction errors and an aggregation bias. Given a specific model, reducing the aggregate prediction error by attempting to reduce the aggregation bias is possible.

In Table 10, the results of aggregate predictions are given for travelers between the same sets of specific zone pairs as were used for the work summarized in Table 9, but now it is stratified into those who are car owners and those who are not car owners. This sort of stratification is quite common in travel demand forecasting (15), and it is evident from the analysis of the differences in the errors of the predicted shares of car between Son and Breugel and Best that such a stratification scheme would improve the aggregate predictions. The improvement in the aggregate predictions that is due to stratification by car availability is shown by the data given in Table 11. Further stratification by choice set (such as by moped availability) could be expected to lead to improved aggregate predictions. The results given in Tables 10 and 11 can be considered acceptable in view of the coarse zoning system adopted and the consequential effects of this on the values of the level-of-service variables applied. Furthermore, for each individual group, there is also an error in the observed share, relative to the total population, that is due to sampling. Because this error increases with decreasing sample size, the greater the number of the observed travelers is that is used to compute the observed aggregate share, the more meaningful the aggregate prediction tests are. Thus an adequate evaluation of the aggregate prediction errors can be undertaken only for those zone pairs with a high trip density, that is, those where the error in the observed share could be assumed to be negligible.

The disaggregate prediction tests have demonstrated that model 14 reproduces the average choice of individuals satisfactorily. The aggregate predictions have much more important implications with respect to the usefulness of the models because the model is tested in the same way as the way in which it will be applied generally. The aggregate modal-choice predictions based on a stratification of travelers into those with a car available and those without a car available can be considered satisfactory given the limitations of the data used for this test. This implies that the model can be applied usefully to aggregate predictions in transportation planning studies.

Table 10. Aggregate prediction results with market stratification by car availability.

Car Available	Sample	Sample Size	Type of Data	Percentage of Shares: Car	Bicycle	Moped	Bus	Train
Yes	Best to Eindhoven, zone 2	30	Predicted	81.67	4.47	11.10	1.73	1.03
			Observed	63.33	13.33	23.33	0	0
	Best to Eindhoven, zone 3	29	Predicted	70.93	3.07	15.59	5.83	4.59
			Observed	62.07	3.45	6.90	17.24	10.34
	Son and Breugel to Eindhoven, zone 2	30	Predicted	74.10	6.00	18.50	1.07	0
			Observed	76.67	20.00	3.33	0	0
	Son and Bruegel to Eindhoven, zone 3	21	Predicted	77.24	4.71	16.71	1.29	0
			Observed	76.19	4.76	14.29	4.76	0
No	Best to Eindhoven, zone 2	16	Predicted	0	22.69	68.81	5.56	2.88
			Observed	0	50.00	43.75	0	6.25
	Best to Eindhoven, zone 3	11	Predicted	0	12.00	35.36	21.09	31.54
			Observed	0	0	36.36	18.18	45.45

Table 11. Aggregate prediction results with and without market stratification.

Sample	Sample Size	Type of Data	Percentage of Shares: Car	Bicycle	Moped	Bus	Train
Best to Eindhoven, zone 2	46	Prediction without stratification	66.85	8.43	20.37	2.78	1.59
		Prediction with stratification	53.26	10.81	31.17	3.06	1.67
		Observed	41.30	26.09	30.43	0	2.17
Best to Eindhoven, zone 3	40	Prediction without stratification	60.35	4.48	17.83	8.98	8.45
		Prediction with stratification	51.42	5.53	21.03	10.03	12.00
		Observed	45.00	2.50	15.00	17.50	20.00

CONCLUSIONS

The modal-choice model for work trips described in this paper was probably the first attempt to consider the full variety of travel modes available in medium- and small-sized Dutch communities where the conventional binary choice model for car and public transit that commonly is used is clearly not suitable. Although the models developed require further work before they could be considered standard operational production techniques, existing models could serve usefully in various transportation planning studies.

A number of conclusions can be drawn from the models estimated. One is that the probability that anyone will choose a given mode is determined largely by factors other than the level of service offered by that mode. If a car is perfectly available to a traveler, then there is a very high probability that he or she will choose it regardless of the characteristics of the alternative modes. The policy implications of this as proposals for traffic restraint in urban areas increase, at least in Europe, are considerable. The estimation results tend to confirm the general assumptions about the relative weights of IVTT and OVTT, although it would appear that there could be significant differences in the evaluation of different types of out-of-vehicle travel time. IVTT would, on the contrary, seem to be viewed similarly for all modes. Travel costs do not significantly influence modal choice for the particular data set, nor do socioeconomic characteristics other than vehicle availability.

It has been suggested previously that the estimation of disaggregate models requires fewer observations than does the calibration of aggregate models. Estimating the same model with different data sets tends to confirm this belief in that the marginal value of increasing the sample size above some 300 observations was found to be small.

The transformation of disaggregate models to aggregate models for use as predictive models presents a number of theoretical problems. It would seem possible, however, that, for all practical purposes, the effects of the problems can be minimized by use of market stratification.

ACKNOWLEDGMENTS

This paper is based on the results of a study carried out by Buro Goudappel en Coffeng b.v., and Cambridge Systematics, Inc., under contract to Projectbureau Integrale Verkeers- en Vervoerstudies, Ministerie van Verkeer en Waterstaat, Netherlands. The permission of the Projectbureau to utilize the results of that study is gratefully acknowledged. However, the opinions expressed are ours and not necessarily those of the Projectbureau. We wish to acknowledge the advice and encouragement of the 2 project monitors, Chresten J. Steilberg and Klaus Broersma, of the Projectbureau. The assistance of Koos Mars, Jan Jetten, and Meindert Bovens in the execution of the study also is acknowledged.

The material in this paper is drawn from a report that is available elsewhere (4).

REFERENCES

1. M. E. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, Massachusetts Institute of Technology, PhD dissertation, 1973.
2. VODAR, Huisenquete: Ontwerp en uitvoering. Buro Goudappel en Coffeng, b.v., Netherlands, 1971.
3. Disaggregate and Simultaneous Travel Demand Models: A Dutch Case Study. Buro Goudappel en Coffeng, b.v., and Cambridge Systematics, Inc., 1974.
4. M. G. Richards and M. E. Ben-Akiva. A Disaggregate Travel Demand Model. Saxon House Studies, D. C. Heath, Ltd., England, 1975.
5. Statistiek van het personenvervoer 1971. Centraal Bureau voor de Statistiek, Staatsuitgeverij, The Hague, Netherlands, 1973.
6. A Disaggregate Behavioral Model of Urban Travel Demand. Charles River Associates, 1972.
7. De vervoerswijzen van de inwoners van het gebied van onderzoek. Commissie Bevordering Openbaar Vervoer Westen des Lands, The Hague, Netherlands, 1970.
8. F. X. de Donnea. The Determinants of Transport Mode Choice in Dutch Cities. Rotterdam Univ. Press, Netherlands, 1971.
9. P. T. McIntosh and D. A. Quarmby. Generalised Costs and the Estimation of Movement Costs and Benefits in Transport Planning. Department of the Environment, London, MAU Note 179, 1970.
10. Implementation of the n-Dimensional Logit Model. Peat, Marwick, Mitchell and Co., 1972.
11. R. H. Pratt and T. B. Deen. Estimation of Sub-Modal Split Within the Transit Mode. Highway Research Record 205, 1967, pp. 20-30.
12. M. G. Richards, B. E. N. Verberne, and G. L. S. Willemse. VODAR: Some Aspects of the Home Interview Survey. Traffic Engineering and Control, 1972.
13. M. G. Richards and M. E. Ben-Akiva. A Simultaneous Destination and Mode Choice Model for Shopping Trips. Transportation, Vol. 3, No. 4, 1974.
14. P. R. Stopher, B. D. Spear, and P. O. Sucher. Toward the Development of Measures of Convenience for Travel Modes. Highway Research Record 527, 1974, pp. 16-32.
15. A. G. Wilson. Entropy Maximizing Models in the Theory of Trip Distribution. Journal of Transport Economics and Policy, Vol. 3, No. 1, 1969.

STRUCTURAL TRAVEL DEMAND MODELS: AN INTERCITY APPLICATION

John B. Peers and Michael Bevilacqua, Alan M. Voorhees and Associates, Inc.

Conventional sequential transportation models clearly have limitations as estimators of intercity travel demand. Despite their theoretical advantage, little work has been carried out in the full application of behavioral or "structural" models. Structural-model development is focused primarily on disaggregate models, particularly for modal split. This paper discusses the development of an alternative approach, that of developing a set of direct-demand models for estimating intercity transit travel for a Sacramento-Stockton-San Francisco Bay Area corridor study. A series of judgments are described that identify why structural models rather than sequential models were chosen and why direct-demand models rather than probabalistic-choice models were used. The methodology of calibratión, including variable selection and equation development, validation, and forecasting, is outlined. Emphasis is placed on the trade-offs to be made among policy responsiveness, accuracy, and the practical problems of developing and using such forecasting tools. The material has been oriented toward the planner-engineer faced with the practical issues of selecting and using intercity travel demand forecasting procedures.

•ENGINEERS and planners in transportation forecasting have become more aware recently of the changing and searching questions that they are required to answer. They also are aware that existing modeling techniques, particularly the best known models forming the sequential decision-making process, have severe shortcomings in their abilities to answer these questions (1). In the planning of the 1950s and 1960s, the emphasis was on building new transportation facilities, which were nearly always highways, to maintain or improve existing levels of service and to match a long-term demand forecast. The major restraint on such plans was the budget. Large quantities of money, particularly the Highway Trust Fund, were set aside for rural and urban freeway facilities. The physical structure had been anticipated, and concern was on the size of structures in terms of the number of freeway lanes and capacities at intersections.

For a transportation corridor study between Sacramento, Stockton, and the San Francisco Bay Area, concerns were with the development of a staged plan for transportation (specifically transit) improvement. Putting all the findings in a format that could be understood by many people rather than precisely understood by a few also was necessary. The clients for the study, the California Department of Transportation, the California State Senate, and the U.S. Department of Transportation, reinforced the need for general understanding. In addition, the clients wanted to know what kind of assumptions were included when patronage estimates were made, whether the assumptions (such as parking or fare costs or frequency of service) were open to public policy change, and what effect such changes would have on transit programs in terms of market response of riders and consequent financial costs and revenues.

In such an environment, to combine the benefits of the structural models with the advantages of logical behavioral relationships, responsiveness to differing assumptions of policy issues, and speed of turnaround when questions arose that required additional analysis clearly was mandated.

PROBLEM OF CHOICE-MODELING DECISION

The history of travel forecasting has been one of successively more comprehensive attempts to move from models that simply project demand to those that provide a coherent representation and organization of the complex of consumer attitudes, behavior, and perceptions of service attributes that produce travel demand. The structure of such models should, in theory, permit them to respond to significant changes in the transportation service variables specified for the model regardless of whether the level of service associated with a specific model has been experienced previously.

A primary objective of the Sacramento-Stockton-San Francisco Bay Area corridor study was the assessment of the feasibility of alternative forms of transit systems and the evaluation of their impacts across a wide range of issues. In conjunction with this, the necessity to effectively forecast the possible demand for intercity travel became apparent. For most of the systems proposed for the corridor, no previous operational experience existed within this region from which data on travel characteristics could be monitored and collected. Therefore, if a model that could effectively estimate travel on these "new" systems was to be employed, the model had to be responsive to service as well as to user attributes. The underlying strategy associated with the estimation of the demand for intercity travel was to develop a series of models that established predictable relationships among physical systems, demographic characteristics, activity distributions, and travel behavior. Several specific criteria were defined in the effort to develop a demand-estimation tool that would have

1. The ability to incorporate a broadened range of such service characteristics of the transportation system;
2. The capability of incorporating responsiveness to nontransit events, such as gasoline price increases and speed limits, into its structure;
3. Transferability to other corridors; and
4. Both long-range and short-range usefulness not only as a planning tool but also as a link-specific design tool for new system improvements.

Conventional urban transportation models for estimating travel demand have a number of deficiencies that limit their validity and utility in estimating both travel behavior and patterns and the impact assessment of new transportation facilities and modes.

1. The estimated number of trips produced by a household is typically not sensitive to the quality of service provided by the transportation system. Accordingly, conventional models show travel demand as being insensitive to service whether there are 1, 2, or 3 transportation modes available and whether the transport facilities available are continually overloaded or are continually free-flowing. As a consequence, no direct mechanism exists in the demand-estimation process to deal with latent or induced travel demand or to reflect transportation system quality or transportation pricing effects.
2. Most conventional models are sequential and involve 4 step functions in estimating demand: trip generation, trip distribution, modal choice, and route assignment. The sequence in which these functions are carried out predetermines the underlying behavioral rationale and presupposes a travel-decision process that is not substantiated with factual behavioral research.
3. Too many conventional models are derived from empirical data-fitting without taking account of any underlying theoretical foundations or behavioral hypotheses. Consequently, their behavioral properties are suspect and their utility for travel-demand forecasting or policy analysis under differing conditions and constraints is highly unsatisfactory.
4. Many conventional models lack any direct expression of public policy variables in their formulation. As a result, their use and value in planning analysis are restricted substantially.
5. The lack of fundamental theory and behavioral properties underlying most conventional models, as well as their failure to incorporate policy sensitivity, makes the

transferability of most models highly dubious. This means that the model relationships developed for that area can seldom be translated for use in other urban areas; it also means that the issue of "new mode" or facilities introduced into an urban- or intercity-corridor setting poses a serious problem for conventional models because of their behavioral deficiencies and questionable forecasting reliability. The record to date in widely varying patronage projections for the new system is ample evidence of this problem.

The first 4 of these limitations could not be readily accepted for the intercity-corridor study. Such features as the identification of causal relationships between trip making and user and system attributes and the ability to express the decision to travel as a simultaneous function of mode, destination, and route made it clear that structural models should be selected as the most effective technique for satisfying intercity-travel requirements.

The terms behavioral and structural are commonly interchanged freely in modeling. Structural models that can be specified so that they relate the decision to travel to the characteristics of the trip maker can be considered behavioral (2).

STRUCTURAL MODEL ALTERNATIVES

Structural models can be separated into 2 distinct classes: direct demand and probabilistic choice. Direct-demand models estimate travel demand by origin, destination, and mode with a single equation (3). Probabilistic-choice models estimate the probability of choosing 1 alternative from a set of available alternatives (4). More specifically, the probabilistic-choice model potentially estimates the probability or likelihood of a traveler's making a trip conditional on 6 decisions—frequency of trip, destination, mode, time of day, choice of route, and purpose—or on a subset of these 6 decisions. These probabilities are evaluated on a per-person or per-household level. To determine the absolute number of trips within each category, one must multiply this function by the total number of households or persons at the origin zone.

The decision to choose the aggregate direct-demand model to estimate future travel demand in the Sacramento-Stockton-San Francisco Bay Area corridor was based on several issues including the availability and requirements for data, experience of model use, special features of each model, and subsequent costs in time and money.

Data requirements perhaps constitute the foremost constraint to the development of a probabilistic-choice model. Both direct-demand and probabilistic-choice models can be calibrated with either aggregate or disaggregate data. It is accepted that modal split, or market share, is a function of socioeconomic indicators such as income. It follows that probabilistic-choice models, which define market share, respond best to market-segmented, or disaggregate, data. This presents several problems. Existing travel information as compiled by the origin-destination surveys conducted in the 1960s generally is not in a format that is compatible with the calibration of disaggregate models. Therefore, expending significant efforts to reformat the data becomes necessary. In most cases, and in this study, the time required for the compilation of base-year household data to obtain information on choice of mode, frequency of travel, choice of destination, time of day, choice of route, and purpose precluded pursuing this course. In the corridor study, the option of using market segmentation that stratified the data by income class, household-ownership category, and household size was considered to minimize base-year data reformatting; however, it would have been necessary to calibrate 90 models for the transit mode alone (5 purposes × 3 household sizes × 3 income classes × 2 household-ownership categories). In contrast to this, it was estimated that only 5 direct-demand models for transit need be calibrated.

In addition to base-year data format needs and the number of models requiring calibration, there is the issue of aggregate or disaggregate modeling in the future year. Estimating future-year population and employment to produce reasonable and reliable results is difficult. Reliable techniques have not been devised by which reasonable estimates can be expected for substratifications of population and employment. To obtain such figures would require the application of extrapolated "factors" that, in turn, are highly vulnerable to error. The suggestion is that, the higher the level of disaggregation is, the less reliable the data become.

Experience and competition by destination and mode also were important issues. With regard to experience, an additional disadvantage associated with the development of probabilistic-choice models is that there is practically no production experience in their development and application as predictive models. Although a maximum-likelihood technique for the calibration of probabilistic-choice models has been developed, there have been limited opportunities to apply and test the results of this procedure. Until recently, most probabilistic-choice models, developed as operational rather than research tools, have been used as modal-split or explanatory models (5).

An advantage of the probabilistic-choice model is its sensitivity to competing activities and competing systems. The problem of competing activities has been partially overcome by the proper specification of the direct-demand model, that is, by the expression of the attraction variables in a form that represents the market share as opposed to the magnitude. The ability of direct-demand models to respond to alternative-mode system changes depends to a large extent on the ability to include a comprehensive set of alternative-mode system variables in the model. This could be achieved for transit models in the corridor study.

Having considered the problems associated with assembling base-year disaggregate data, the significantly increased effort implied by calibrating and estimating 90 models, the forecasting of market-segmented data, and the inadequate production-oriented experience of probabalistic-choice models, we decided that an aggregate direct-demand modeling procedure would be the most feasible approach to pursue. We decided to calibrate less precise models with good forecast data rather than to define highly refined models with forecast data of questionable reliability.

DIRECT-DEMAND MODELS

Direct-demand models can be specified as either modal-abstract models or modal-specific models. The primary advantage of a modal-abstract model is that only 1 equation is necessary to estimate travel demand (6). This is particularly advantageous when one is estimating demands for new modes that are not in operation or for which there are no existing prototypes. The primary disadvantage associated with developing a modal-abstract model, however, is that it requires that each alternative mode be described by a single set of variables. The selection of a set of attributes that can effectively represent the wide range of system features characterizing different modes can present a major problem because homogenizing attributes means the loss of model responsiveness to policy changes. Further, the ability to identify cross elasticities becomes impaired. An attempt was made, however, to calibrate a set of modal-abstract models; it was unsuccessful because of data inadequacies.

Modal-specific models require separate formulation of generically different modal forms. Although it may be possible to develop separate models for automobile, bus, rail, and airplane modes, the distinction generally is limited to automobile and transit. The separate formulation of models by mode provides the opportunity to achieve the maximum flexibility in model specification. Furthermore, by modal-specific modeling, those intrinsic qualities associated with the automobile, such as privacy and convenience, as well as those associated with transit, will be reflected in the model calibration. Given the provision for greater system and user sensitivity that is afforded by modal-specific models, we decided to adopt these functional forms (that is, automobile and transit) in the development of the direct-demand models.

It was necessary to determine which model form would be best suited for the intercity application. Various mathematical forms have been suggested and applied in the development of previous demand models. Basically, there are 3 forms (with respect to the variables) that the model function can assume: linear, nonlinear, and mixed. The nonlinear form includes product forms of powers and exponentials. The decision to choose one form over the other is more pragmatic than theoretical. Relatively little research has been conducted to assess the influences of model form on the performance of the model. However, several observations can be made with respect to the most reasonable form for a model. From an examination of the function one can see that the dependent variable in nonlinear models is much more responsive to a given change in an independent variable than it is in linear models. In addition, travel data do not support the idea that trip makers behave in a manner that is responsive to changes in individual causal variables that have been arranged in a linear function. Finally, nonlinear functions have been shown to be more effective in describing observed trip-making behavior. Although the issue cannot be definitively resolved, we found that a mixed-form model provided the necessary versatility and fitted the survey data best.

MODEL CALIBRATION

Before calibration could begin, the data for calibration had to be developed. The prime source of trip data was the Bay Area Transportation Study (BATS) files, developed from home-interview surveys taken in 1965. The data required reformatting into zone data a 123-zone system covering the San Francisco Bay Area, Sacramento, and Stockton. Ninety zones were included in the Bay Area. For the base year, the highway network was derived from California Department of Transportation data. Four transit networks were coded for the base transit system, 1 pair for off-peak travel and 1 pair for peak travel. Each pair had 1 network for public transit access and walk connectors to the transit facilities. The second network had only private automobile access to the transit facilities. Production-zone socioeconomic data related to households were developed by using expanded home-interview data in the Bay Area from BATS and in the Sacramento and Stockton areas from the equivalent home-interview data collected in 1967 and 1968. The attraction-zone socioeconomic data related to employment categories were developed from surveys that were conducted in conjunction with the home-interview surveys.

The data available for use in calibration included 14 household statistics; 3 automobile-service statistics (walk time, in-vehicle time, total cost per car trip); 8 transit-mode statistics for automobile access by submode; 7 transit-mode statistics for nonautomobile access by submode; 10 destination-zone statistics (mainly subsets of employment data); and 4 subsets of trip data for the classes of purpose (home-based work, shop, other, and non-home-based trips).

The first step of the calibration process was that of attempting to find and identify causal variables. Sample statistics of zone means were listed, and a correlation matrix of variables, including the log and exponential forms of the variables, was developed for identifying correlations between independent variables and trips, the dependent variable. Each variable also was plotted against trips. The matrix and plots were reviewed to produce the best set of variables in their best forms for explaining the variance of trip making. In addition, it was necessary to check for levels of independence or low correlation between independent variables. Constraints were applied to some variables, for example, to a relationship between in-vehicle transit time and out-of-vehicle transit time. This was required because the path-builder algorithm requires weighted values of out-of-vehicle travel time to calculate the minimum time paths. The constraint applied to the model variables, therefore, maintained this minimum-path practice by replacing the 2 variables in-vehicle and out-of-vehicle time by 1 variable:

$$QLT + 2.5\ QXT \tag{1}$$

where

QLT = transit line time, and
QXT = transit out-of-vehicle time.

INITIAL MODEL SELECTION

Because an array of variables was known that satisfied the 2 criteria of (a) being strongly correlated with the dependent variable, trip making or demand, and (b) maintaining orthogonality among the exogenous variables, a series of models was produced by using these variables and a nonlinear regression program specifically adopted for this study. The primary virtue of using a nonlinear regression program for estimating the models was that it obviated the need to transform the dependent variables into a linear form. This process of using linear transformations, which is a requirement when applying standard linear regression programs, introduces bias in parameter estimation. The application of a nonlinear regression means that techniques such as restraining variables with reasonable limits need no longer be applied. As a result, the use of nonlinear regression was a significant improvement over standard estimation procedures.

The final set of variables used in specifying the transit models was divided into 3 groups: extensive variables, intensive variables, and system variables. The extensive variables are as follows:

1. Residential population;
2. Employment, by type;
3. Workers; and
4. Locations, magnitude, and net density according to "alternative futures" of moderate northern regional growth with environmental constraints versus slow, southern, dispersed regional growth (current trends).

The intensive variables are as follows:

1. Persons per household,
2. Income per household and income per worker,
3. Cars per household and cars per worker, and
4. Employment per acre (hectometer^2).

The system variables are as follows:

1. Automobile speeds (travel time),
2. Automobile out-of-pocket costs,
3. Transit speed (travel time for feeder and line-haul),
4. Transit costs (for feeder and line-haul),
5. Walking and waiting time,
6. Parking costs,
7. Service frequency (peak and off-peak),
8. Terminals per station locations, and
9. Mode and service path.

For each set of models, 4 major statistics were developed that compared the synthesized trips with surveyed data:

1. Error mean,
2. Absolute error mean,
3. Error mean squared, and
4. Coefficient of determination r^2.

It was important to ensure that all the variables open to policy action and variation, such as parking pricing or fare structure, were included wherever possible in the models. In some cases this meant accepting one model form over a better fitting model because the better fitting model did not include these important variables. Many techniques were used to analyze and compare the different models produced by this process. However, the most important single criterion was judgment. Because the models selected had to make sense, relationships implied by their structure and parameters had to be reasonable.

An initial set of models was calibrated by using a sample set of data taken from half of the surveyed information. A second set of models was subsequently calibrated for the entire surveyed data. Both sets of models were nearly identical. The 5 transit models developed covered the following:

1. Home-to-work trips by transit,
2. Home-to-work trips by automobile,
3. Home-to-shop trips,
4. Other home-based trips, and
5. Non-home-based trips.

The generic form of all the models was:

$$(\pi_1 P_1 + \pi_2 P_2)\,(\alpha_1 A_1 + \alpha_2 A_2)\; Z_1^{\xi_1} Z_2^{\xi_2} X_1^{\Theta_1} X_2^{\Theta_2} e^{\phi_1 X_1} e^{\phi_2 X_2}$$

where

P_1, P_2, A_1, and A_2 = extensive production and attraction variables describing zone size (such as zone populations, employment, and workers);
Z_1 and Z_2 = intensive production and attraction variables such as cars per worker and retail jobs per area;
X_1 and X_2 = interchange service variables by mode such as in-vehicle time, out-of-vehicle time, and out-of-pocket costs; and
π, α, ξ, Θ, and ϕ = model parameters.

MODEL VALIDATION

The models had been calibrated entirely on base-year data (1965) for the Bay Area only. Even though the Bay Area includes most of the zones and socioeconomic activity, there was a requirement to validate the synthesized trips against corridor travel. At the beginning of the study, some effort had gone into surveying intercity travel for both highways and transit (bus) modes. No corridor travel data for transit travelers were available from the BATS files because transit travelers traveling outside the Bay Area had been recorded only to the transit terminals—airport, bus terminals, and railroad stations.

Interchange pairs along the corridor, particularly those with one end outside the BATS area, were compared for synthesized trips from 1965 socioeconomic data and the 1973 surveys plus some data from the Greyhound Bus Company files and the California Division of Highways annual vehicle counts. Trip-length distributions for survey and synthesized trips also were compared. When we amended the constant for each model for each 5-min interval of weighted trip length, the synthesized trips and surveyed trips maintained close relationships for both trip-length-frequency curves and specific corridor interchanges. Finally a stepwise approach was taken to forecast trips by using future socioeconomic data and future transit and highway networks. Initially, the 1995 networks were used together with 1965 socioeconomic data to produce trip tables. Total trips plus major interchanges along the corridors were inspected to see the effect of the presence of an upgraded set of transit networks. Then the 1965 networks were used together with 1995 socioeconomic data to produce trip tables. Again the total trips and the corridor movements were inspected. Two problems became clear from these analyses. One related to maintaining as linear the extensive variables in the models; the other related to large increases in trips due to increased income. The extensive variables defining zone population, employment, and subsets of population and employment always were kept linear; that is, they were kept in a power-product form without exponents other than unity. The maintenance of this linearity was an important condition because the models were independent of zone size and potentially were more transferable to different zone sizes within the study area or elsewhere. The result of this decision, together with the inclusion of both production (population) and attraction (employment) variables in the model, was that, where, for example, both population and employment doubled, the total number of trips increased 4-fold. To overcome this problem, the attraction variables had to be normalized in each model. This was accomplished by replacing the extensive attraction variable by a new variable that reflected the relative increase in the attraction activity. As a result, the models also became sensitive to the notion of market share; that is, if a large number of attractions were to be added to one zone, then the market share of attractions would increase for that zone and the interchanges between the origin zone and that zone would increase relative to the unchanged zones.

The income issue was a problem mainly because the future income was forecast to increase substantially. Average incomes per household at zone levels in 1965 ranged from $4,700 to $13,800. For 1995, estimated average incomes measured in real terms ranged from $11,700 to $35,000. A number of the models were highly elastic with respect to increases in income. Work trips for automobile access and shopping trips, for example, had exponents 2.0 and 1.9. That the models become steadily more unstable as the data stray farther from the base-year ranges is accepted. For future data, constraining income values to reduce the effect of this potentially explosive variable was necessary.

In summary, the calibration, validation, and forecasting of the models were developed in 4 steps:

1. Development of calibration data for trips, socioeconomic data, and networks;
2. Development of equations;
3. Validation against corridor movement; and
4. Cautious manipulation of the models to produce future forecast trip tables.

Each step took considerable levels of both time and effort, but for the transit models each step was carried out successfully.

MODEL APPLICATION

On completion of the calibration and validation stages, we produced a final set of transit models. The specifications of these models are given in Tables 1 through 5. The dependent variable in all of these models is 1-way transit trips. The coefficients of determination and forms of the models are given in Table 6. K is the constant in all of the forms. All of the models yielded reasonable r^2 values. The work-purpose models were the best correlated models with r^2 values of 0.65 and 0.72 for the public-transportation-access and automobile-access models respectively. The remaining purposes had lower correlations. However, in terms of total transit trip making, the effective r^2 value is better than these values might imply because of the dominance of work trips. A weighted average of the r^2 by purpose and percentage of intercity transit trips by purpose will yield an effective r^2 value of 0.64 for the total trip demand.

The transit and highway networks that would be employed in estimating future-year travel demand were developed at the same time as the transit models were developed. Seven distinct transit system alternatives were chosen to be analyzed. System technologies included express bus and conventional and high-speed rail options. For each of these systems, networks representing each of 3 access modes and 2 time periods were constructed. In all, 42 future-year transit networks were built. In addition, 1 future-year highway network was built.

With regard to the application of the transit models, approximately 84 distinct program packages were defined. These program packages were derived from combinations of the system line-haul alternative, the access mode, the line-haul fare, the access fare, and the demographic growth alternative. Additional program packages were derived from combinations of these independent corridor-specific program packages

Table 1. Model for home-based work trips, public transportation access.

Variable	Description	Coefficient a_y
X_1	Automobile out-of-pocket cost/transit fare	0.188
X_2	Income/worker at zone of origin	0.564
X_3	Vehicles/worker at zone of origin	-1.494
X_4	Transit line time/(2.5 + transit wait time)	-1.355
X_5	Transit line time - automobile line time	-0.028
X_6	Automobile out-of-vehicle time	0.275
X_7	Workers at zone of origin	—
X_8	Jobs at zone of destination	—

Table 2. Model for home-based work trips, park-and-ride.

Variable	Description	Coefficient a_y
X_1	Transit fare	-1.137
X_2	Automobile out-of-pocket cost	0.838
X_3	Income/household at zone of origin	2.073
X_4	Vehicles/household at zone of origin	3.864
X_5	Transit wait time + (transit line time/2.5)	-1.161
X_6	Automobile line time	0.401
X_7	Automobile out-of-vehicle time	1.806
X_8	Workers at zone of origin	—
X_9	Jobs at zone of destination	—

Table 3. Model for home-based shopping trips.

Variable	Description	Coefficient a_y
X_1	Transit fare - automobile out-of-pocket cost	-0.0075
X_2	Transit wait time - automobile out-of-vehicle time	0.0316
X_3	Income/household at zone of origin	1.868
X_4	Persons/household at zone of origin	3.839
X_5	Retail jobs/acre at zone of origin	0.233
X_6	Vehicles/household at zone of origin	-2.479
X_7	Transit line time + (2.5 × transit wait time)	-1.619
X_8	Adults at zone of origin	—
X_9	Retail job at zone of destination	—

Note: 1 job/acre = 2.50 jobs/hm².

Table 4. Model for home-based other trips.

Variable	Description	Coefficient a_y
X_1	Transit fare	-0.231
X_2	Income/household at zone of origin	-0.045
X_3	Automobile out-of-pocket cost - transit fare	0.0046
X_4	Automobile line time	0.0439
X_5	Automobile out-of-vehicle time	0.0189
X_6	Transit line time + (2.5 × transit wait time)	-0.0238
X_7	Persons/household at zone of origin	3.214
X_8	Vehicles/household at zone of origin	-1.537
X_9	Transit line time/(2.5 + transit wait time)	-2.27
X_{10}	Population at zone of destination	0.00015
X_{11}	Service jobs at zone of destination	0.00024
X_{12}	Population at zone of origin	—

Table 5. Model for non-home-based trips.

Variable	Description	Coefficient a_y
X_1	Transit fare	-1.352
X_2	Automobile out-of-pocket cost	1.403
X_3	Transit line time/(2.5 + transit wait time)	-4.00
X_4	Automobile line time	0.951
X_5	Automobile out-of-vehicle time	0.184
X_6	Population at zone of origin	0.00044
X_7	Jobs at zone of origin	0.0002
X_8	Population at zone of destination	0.00044
X_9	Jobs at zone of destination	0.0002

(such as a rail system between the Bay Area and Sacramento combined with a bus system between the Bay Area and Stockton). Together, more than 100 program packages were described. Trip tables subsequently were estimated for approximately 40 selected program packages. The analyzed program packages were selected so that intercity travel demand for the remaining alternatives could be estimated by interpolation if desired.

As a final evaluation of the transit models, the 1995 rates for interzone transit-demand generation were compared to the rates observed in 1965. In both cases, the total number of interzone transit trips was compared to the area population. In 1965, the transit trip generation rate was 0.09 trip/person; in 1995, for the high-growth alternative, the rate increased to 0.12 to 0.16 trip/person depending on the program package analyzed. Because the interzone transit-trip totals are biased toward longer trips (shorter, intrazone trips are excluded), to expect generation rate to increase with the improvement of intercity transit is not unreasonable. The stability of these generation rates further substantiated the validity of the models and their demand estimates.

RESPONSE TO ALTERNATIVE POLICIES

Sensitivity analysis is an important result of model development. The capability of the direct-demand models to respond accurately and quickly to alternative assumptions regarding the system and the user, and, therefore, to enable policymakers to see the effect of the policy alternatives they suggest, is a powerful feature. The response of the model to changes can be assessed by the analysis of the elasticities of the transit travel demand with respect to the components of the model. Elasticity can be defined as a dimensionless number that represents the percentage of change in the travel demand for a 1 percent change in any of the independent variables. In defining the elasticity, only 1 variable is changed and the others remain constant. By applying the concept of elasticities, we could analyze the sensitivity of the demand to ranges of the values of the system and user inputs. This technique is useful in analyzing the impact of various policy changes, such as increased fuel prices, decreased automobile speed, and increased transit service frequency on demand for transit travel.

Table 7 gives the elasticities derived for the household variables. Values for the system elasticities have not been shown because they are not always constant. In many cases, they are complex functions and have a meaning only within the context of a specific interchange movement.

As a result of the sensitivity analysis, several interesting relationships can be identified and generalized for the total intercity travel demand in the region.

1. The 30-year increase in corridor population between 1965 and 1995 represented by the low-growth alternative is equivalent to a 60 percent population increase causing a 60 percent increase in total regional transit demand. The moderate-growth alternative is equivalent to an 80 percent population increase causing an 80 percent increase in demand.
2. A 25 percent increase in income per household causes a 26 percent increase in transit demand for intercity travel.
3. A 25 percent increase in car ownership per household causes a 20 percent decrease in total transit trips with transit access, but a doubling of those transit trips with automobile access (mostly long trips).
4. A 100 percent increase in the current price of gas representing a 50 percent increase in out-of-pocket operating costs for automobile travel causes a 60 percent increase in total transit demand for long trips. A 200 percent change in the current downtown parking charge will have the same effect for downtown-oriented trips.
5. For long trips with a 40-min wait and transfer time (assuming that the 40 min is made up of 20 min of walk time and 20 min of wait time), a 50 percent reduction in headway will cause a 60 percent increase in demand. A 10-min reduction in wait time for long trips would produce a similar increase in demand. For shorter trips, this effect would be halved.

Table 6. Coefficients of determination and forms of models in Tables 1 through 5.

Model	Table	r^2	Form
Home-based work, public access	1	0.65	$Trips_{ij} = KX_1^{a_1}X_2^{a_2}X_3^{a_3}X_4^{a_4}e^{a_5X_5}X_6^{a_6}X_7X_8$
Home-based work, park-and-ride	2	0.72	$Trips_{ij} = KX_1^{a_1}X_2^{a_2}X_3^{a_3}X_4^{a_4}X_5^{a_5}X_6^{a_6}X_7^{a_7}X_8X_9$
Home-based shop	3	0.43	$Trips_{ij} = Ke^{(a_1X_1-a_2X_2)}X_3^{a_3}X_4^{a_4}X_5^{a_5}X_6^{a_6}X_7^{a_7}X_8X_9$
Home-based other	4	0.47	$Trips_{ij} = KX_1^{a_1}X_2^{a_2}e^{a_3X_3}e^{a_4X_4}e^{a_5X_5}e^{a_6X_6}X_7^{a_7}X_8^{a_8}X_9^{a_9}(a_{10}X_{10} + a_{11}X_{11})X_{12}$
Non-home-based	5	0.54	$Trips_{ij} = KX_1^{a_1}X_2^{a_2}X_3^{a_3}X_4^{a_4}X_5^{a_5}(a_6X_6 + a_7X_7)(a_8X_8 + a_9X_9)$

Table 7. Elasticities for household variables.

Trip Purpose	Persons/ Household	Income/ Household	Vehicles/ Household
Work, public access	N.A.	0.56[a]	-1.49[b]
Work, park-and-ride	N.A.	2.07	3.86
Shop	3.84	1.87	-2.48
Other	3.21	-0.05	-1.54

Note: N.A. = not applicable.

[a]Income/worker. [b]Vehicles/worker.

6. For long trips (more than 1 hour), an automobile speed limit decrease from 65 to 55 mph (104 to 88 km/h) will cause a 45 percent increase in total transit demand. For short trips, the same change in automobile travel time will result in a 14 percent increase in transit demand.
7. A 50 percent decrease in transit block time, such as the difference between track-levitated vehicle and turbotrain between Oakland and Sacramento, will cause a 200 percent increase in transit demand.
8. An increase in total fares (access and line-haul) for a long trip, such as from San Francisco to Sacramento, from $5.00 to $10.00 will cause a 40 percent decrease in transit demand.

SUMMARY

In the past few years, awareness of the limitations associated with the application of conventional sequential models (generation, distribution, and modal-split models) has been increasing. The structural models that have been recommended as replacements have covered a wide variety of model forms and calibration processes. Yet, despite the large number of alternative modeling choices made available, few studies attempted to use other than sequential models.

We feel that, at least in part, the lack of acceptance of structural models stems from a lack of a basic understanding of the features and applications of these models. For the intercity-corridor study in Northern California, we have found that a power-product, aggregate direct-demand model most successfully satisfied the objectives that were developed early in the study. These objectives included policy sensitivity and demand response to alternative transportation systems.

In the calibration of the direct-demand model, we found that the use of a nonlinear regression technique overcame many of the problems of variable transformations and constraints that have been encountered in previous studies. In addition, however, application of the model equations in the future still will require the careful, judgmental processes used in the sequential models. Particular attention needs to be paid when one attempts to estimate future demand by using future socioeconomic and system data that are outside the range of the base-year data.

Finally, the results of the modeling work provided an opportunity for a clear and useful dialogue between the technicians and the policymakers. As a result, the policymakers were afforded a technique by which they could test and assess the effects of a variety of alternate policy assumptions.

ACKNOWLEDGMENTS

The models presented in this paper were developed by Alan M. Voorhees and Associates, Inc., as part of a study sponsored by the California Department of Transportation.

We wish to thank Daniel Brand of Harvard University and James McLynn of DTM Inc. for their assistance and suggestions during the course of the model development. We also wish to extend our thanks to Jeffrey MacMann for his hard work and invaluable efforts throughout this phase of the study. All errors, however, are our sole responsibility.

REFERENCES

1. D. Brand. Theory and Method in Land Use and Travel Forecasting. Department of City Planning, Harvard Univ., 1972.
2. T. J. Adler and J. A. Bottom. Formulation of Travel Demand Modelling Requirements. Department of Civil Engineering, Massachusetts Institute of Technology, 1973.
3. T. A. Domencich, G. Kraft, and J.-P. Valette. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968, pp. 64-78.
4. D. Brand. Travel Demand Forecasting: Some Foundations and a Review. Department of City Planning, Harvard Univ., 1972.
5. M. E. Ben-Akiva. Passenger Transportation Demand Theory and Models. Department of Civil Engineering, Massachusetts Institute of Technology, 1972.
6. E. R. Ruiter. Analytical Structures. Special Rept. 143, 1972, pp. 178-205.

JOINT-CHOICE MODEL FOR FREQUENCY, DESTINATION, AND TRAVEL MODE FOR SHOPPING TRIPS

Thomas J. Adler and Moshe Ben-Akiva, Department of Civil Engineering, Massachusetts Institute of Technology

This paper describes the estimation of a disaggregate joint-choice model for frequency, destination, and travel mode for shopping trips. The model builds on earlier research by Ben-Akiva in Transportation Research Record 526 that argued for the replacement of aggregate conditional (or sequential) model systems with disaggregate joint (or simultaneous) models and presented a model for the joint choice of destination and mode for shopping trips. The extension of this general model by use of the same multinomial logit form and a similar specification to include travel-frequency choice is an attempt to provide a more complete version of the joint-model structure. Estimation of the expanded joint-choice model proved to be feasible and resulted in behaviorally and statistically acceptable parameter values. All variables produced coefficients of the expected signs and magnitudes consistent with the behavioral notions on which the model specification was based. In general, the estimation of joint-choice models for travel demand was shown to be a computationally tractable alternative to the less acceptable conditional approaches that have been used in the past. An example of the application of the shopping model (combined with a previously estimated modal-choice model for work trips) to the evaluation of transportation policy options is used to highlight some of the features of both the particular models used and the general modeling approach that they represent.

•THIS PAPER describes a disaggregate travel demand model based on a joint-choice structure. The development of disaggregate travel demand models has led from initial work on binary modal-choice models to continual expansion of the context of choice in an effort to produce a complete set of models for predicting urban travel patterns. Two recent research projects have set the stage for the results reported in this paper. The first project estimated a disaggregate travel demand model for the choices of frequency, destination, mode, and time of day for shopping trips (4). This model was based on the assumption of a conditional-choice (or sequential-choice) structure. Then Ben-Akiva (2, 3), using theoretical arguments, proposed the joint-choice structure as a more realistic approach for travel demand models. His empirical study centered around the choices of destination and mode for shopping trips. Models were estimated for modal choice followed by destination choice, destination choice followed by modal choice, and the joint (or "simultaneous") choice of destination and mode from the set of alternative combinations of mode and destination. As anticipated, empirical evidence has shown that the coefficient estimates are sensitive to the choice structure on which estimation is performed. This finding, along with theoretical arguments, forms a convincing case for the use of a joint-choice structure for all hierarchical equivalent and interdependent choice dimensions such as frequency, mode, destination, and time of day for shopping trips.

The model described in this paper extends the joint-choice model for destination and travel mode estimated by Ben-Akiva to the inclusion of the third choice dimension that is commonly of interest in travel demand forecasting—frequency choice. Consistent with the previous model, the unit of travel demand that is modeled is a round trip (home-shopping-home). The behavioral unit is the household, which is recognized as the relevant decision-making unit for shopping travel choices. As in most of the previous disaggregate choice models with multiple alternatives, the multinomial logit model was used because of its many desirable theoretical and computational advantages over other techniques.

The first part of the paper is devoted to the description of the joint-choice shopping travel demand model, its specification, and the estimation results. [A more complete description of the accompanying research and findings is given elsewhere (1).] Following the sections that describe this model, an example of the application of the shopping model (combined with a previously estimated modal-choice model for work trips) is described. This demonstration highlights some of the important features of both the particular models used and the general modeling philosophy that they represent.

MODEL

This model forecasts the short-term travel choices of a household given predetermined residence location, automobile ownership, and choice of mode to work. By using the multinomial logit form, one can express the joint-choice travel demand model for a given trip purpose as follows:

$$P(f, d, m) = \frac{\exp(V_{fdm})}{\sum\limits_{f'd'm' \in FDM} \exp(V_{f'd'm'})} \tag{1}$$

where

$P(f, d, m)$ = probability of choice of a given frequency, destination, and modal combination;
V_{fdm} = utility of an alternative fdm combination; and
FDM = set of available alternatives where an f, d, m combination represents an alternative trip.

V_{fdm} is a function of the independent variables as follows:

$$V_{fdm} = X'_{fdm}\theta^{fdm} + X'_{f}\theta^{f} + X'_{d}\theta^{d} + X'_{m}\theta^{m} + X'_{fd}\theta^{fd} + X'_{fm}\theta^{fm} + X'_{dm}\theta^{dm} \tag{2}$$

The θs in equation 2 are vectors of coefficients to be estimated, and the Xs are vectors of variables defined as follows:

X_{fdm} = variables that differ among all alternatives,
X_f = variables that differ only among frequencies,
X_d = variables that differ only among destinations,
X_m = variables that differ only among modes,
X_{fd} = variables that differ only among frequencies and destinations,
X_{fm} = variables that differ only among frequencies and modes, and
X_{dm} = variables that differ only among destinations and modes.

The most important variables that can be strongly justified on deductive grounds, for the shopping joint-choice model, fall into 4 of these classes:

X_{fdm} = travel cost (such as time, money, and convenience);
X_f = socioeconomic characteristics of the household (such as household size, life cycles, occupational status, income, and automobile ownership);

X_{fd} = attractiveness of destination to the given trip purpose (such as retail employment and floor area); and
X_{fm} = modal-specific variables (such as availability of the automobile for shopping travel and transit convenience).

This decomposition of the joint utility function is useful for the interpretation of conditional probabilities. For example, the conditional probability of modal choice for given frequency and destination derived from the joint-choice model is:

$$P(m \mid f, d) = \frac{\exp(V_{m \mid fd})}{\sum_{m' \in m_{fd}} \exp(V_{m' \mid fd})} \tag{3}$$

The choice set M_{fd} includes all alternative modes available for the given frequency and destination, and the conditional modal-choice utility is

$$V_{m \mid fd} = X'_{fdm}\theta^{fdm} + X'_{m}\theta^{m} + X'_{fm}\theta^{fm} + X'_{dm}\theta^{dm} \tag{4}$$

The variables X_f, X_d, and X_{fd} have no effect on the conditional modal-choice probability. Similarly, X_d, X_m, and X_{dm} have no effect on the conditional frequency choice, and X_f, X_m, and X_{fm} have no effect on the conditional destination choice. It should be noted, however, that all the variables in the joint utility function affect the marginal choice probabilities for all the 3 dimensions of choice. For example, the marginal probability of frequency choice could be expressed as:

$$P(f) = \frac{\exp(X'_f\theta^f + \ell n\, P^D_f)}{\sum_{i \in F} \exp(X'_i\theta^f + \ell n\, P^D_i)} \tag{5}$$

where

$$P^D_i = \sum_{j \in D_i} \exp(X'_j\theta^d + X'_{ij}\theta^{fd} + \ell n\, P^M_{ij}), \text{ and}$$

$$P^M_{ij} = \sum_{k \in M_{jf}} \exp(X'_{fjk}\theta^{fdm} + X'_k\theta^m + X'_{fk}\theta^{fm} + X'_{jk}\theta^{dm})$$

Thus a change in the value of any variable in the joint-choice model will affect all of the marginal choice probabilities.

The logit formulation not only allows specifications of models, including all of these types of variables, but also permits considerable freedom in the composition of alternative sets (FDM) so that the definition and number of alternatives made available to each individual observation in the sample can be varied. The specification of the model consists of the formulation of the utility functions and the definition of the alternatives in the choice set.

DATA

The data used for model estimation were derived from the Metropolitan Washington (D.C.) Council of Governments (WCOG) 1968 home-interview survey and from additional information compiled by WCOG and R. H. Pratt and Associates. The travel information in the home-interview survey consists of household questionnaire responses detailing the travel activity for a 24-hour period of all household members 5 years old or older. For this modeling effort, the surveyed households first were reduced to 25 percent of the original sample, and then the remaining observations were screened for missing or poorly coded information. This left 4,097 households in the working sample. However, of the households in which some shopping travel was reported over the survey period, not all exhibited the simple home-shop-home behavior that was to be modeled. Of the 1,259 households that reported 1 or more shopping sojourns, 501 had traveled in the simple pattern to be modeled, and 403 of these had used either automobile or bus (the 2 modes to be modeled). Because of the systematic reduction of the sample of households that had made shopping trips, the sample of households reporting no shopping travel also had to be reduced to maintain correct proportions of the 2 types of household. (Sensitivity runs were performed to determine the bias in alternative-specific variables due to incorrect proportions in the estimation sample. Although the bias was significant for extremely nonrandom proportions, it became negligible within 10 to 20 percent of the true proportions.) Thus these were reduced to 910 leaving a total estimation sample size of 1,313 households. [The sample of households making shopping trips was reduced to 32 percent of the original (from 1,259 to 403); therefore, the sample of households not making shopping trips was similarly reduced to 32 percent (2,838 to 910).] The relation of this final estimation sample to the original survey sample of households is shown in Figure 1.

In addition to the travel observations, a major data item to be prepared was level-of-service information not only for the observed travel but also for all alternative shopping trips. These data for highway and transit networks had been compiled previously for the Washington, D.C., area.

ALTERNATIVES AND VARIABLES

The frequency alternatives, as represented in the model, were a choice between a simple home-to-shopping-to-home round trip and the no-travel option. However, one aspect of the no-travel alternative deserves specific note. The original home-interview survey from which the estimation samples were derived included travel information for vehicular trips only (except for the first journey to work for which walk trips were recorded). Thus, the no-travel alternative for the shopping model implicitly includes a walk-to-shopping option. This was a factor that had to be accounted for directly in the derivation of the utility functions for the frequency alternatives.

In modeling frequency choice, it was assumed that 3 sets of effects influenced a household's probability of making a shopping trip. The first set is called here the generating effects, which are those characteristics of the household that would make that household more likely to reach the threshold of need for a shopping sojourn on a given day. One such effect is household size, which would account for the rate of growth of a need for shopping activity as well as for the availability of household members to devote time to the activity. Additional variables to measure the structure of the household, such as the ratio of workers to nonworkers and the life cycle of the household, also were considered to be household generating effects. Similarly, income was hypothesized to represent the ability of a household to stock large quantities of goods and thus avoid the general disutility of travel.

A second group of attributes that is seen to affect frequency of shopping travel is the set of variables that measure the impedances of travel to shopping destinations and the attractiveness, for shopping purposes, of these destinations. The notion that the threshold level of need for shopping travel varies with the costs of travel and with the probability of finding the desired goods (7) corresponds behaviorally with the inclusion of this set of effects. Measurement of the levels of service to the various destinations is relatively straightforward, but determination of attractiveness to the household is less apparent.

The third factor included in frequency choice is a measure to account for the accessibility of shopping destinations by walking trips. This is necessary compensation for the fact that the no-travel alternative includes possible walking trips (these were not recorded in the survey). However, because most travel-forecasting applications deal with vehicle-transportation options, the ability to separate walking trips (which, in the home-interview survey, were defined to include bicycle, motorcycle, and other "miscellaneous" modes) from no travel was not considered important for this model.

The modal-choice alternatives were limited, for this model, to automobile (driver and passengers from the same household) and transit. Other modes and modal combinations accounted for 71 of the original 501 observations of simple shopping trips. Although automobile passengers (with drivers not from the same household) accounted for the largest number of these (42), these were not explicitly modeled because they were interhousehold shared rides for which no information was collected on the number of persons in the automobile or sharing of costs. (All intrahousehold shared rides were explicitly identified and modeled as single shopping trips; transit fares were multiplied by the number of persons making the trip together.) The next largest category was taxicab passenger; however, there were only 7 observations for this category.

The automobile alternative was given only to those households that owned at least 1 automobile; automobile ownership was assumed here to be a predetermined choice for shopping travel choices. The bus transit alternative was given only to those households for which a station was accessible within 0.5 mile (0.8 km) of the residence location (the household's location also was assumed as a predetermined choice). For those households that did have transit accessible, the alternative was allowed only to those destinations that also were served by transit.

The types of variables used to model the modal choice include, of course, generic level-of-service variables and additional variables to account for modal-specific effects. The level-of-service variables in this model are in-vehicle travel time (IVTT), out-of-vehicle travel time (OVTT), and out-of-pocket cost (OPTC). IVTTs for both automobile and bus were taken from networks and computer over the shortest path. OVTTs, however, were measured differently for the 2 modes. For bus, the measured OVTT includes an average walk time to the station, wait time for the bus (a varying percentage of the headway), and additional wait times if transfers are necessary. For automobile, OVTTs were taken from zone vectors supplied by WCOG, which set origin terminal times at an average of 2 min (to allow start-up time and the like) and computed destination terminal times depending on the expected difficulty of finding a parking place near the actual destination.

OPTCs for bus trips were the designated fares (multiplied by the number of persons in the same household making the trip). For automobile trips, costs were in 2 portions. The first was expected parking costs (taken from zone vectors); the second was calculated based on origin-destination (O-D) highway distances and travel times. The travel times were used to compute a cost per mile (kilometer) (which varied according to the computed average speed) for fuel and maintenance costs. This was then multiplied by the travel distances to provide an O-D travel cost.

The only other useful piece of modal information available was the number of bus transfers required, but, because transfer time was included in the computation of OVTTs for the bus, the number-of-transfers variable proved to be insignificant for explanation in models that also included OVTTs.

The selection of an alternative set for destination choice was much less straightforward than for frequency and modal choices. To begin with, all of the available information that was to be used for model estimation was based on the WCOG Transportation Planning Board zone and district boundaries. These data included levels of service to zone destinations as well as figures on zone-based retail employment. Although an attempt might have been made to identify specific activity sites, as has been done elsewhere (6), the time and money required for this task were not available. Therefore, an alternative scheme was developed to use the zone-based data.

The selection of a candidate set of destination alternatives for each household was based on district-level trip matrices for the shopping purpose (134 districts in the metropolitan area). All districts for which at least 1 shopping trip was recorded from the household's residence district were allowed as destination alternatives to that household. The idea for using the trip matrix for this initial assignment of alternatives was to ensure that all destinations with positive probabilities (at least in the estimation sample) were given as alternatives. The district level was chosen both because it was convenient to use with the available level-of-service and attraction data and because it seemed to correspond most closely with a perceptual breakdown of the urban area into shopping opportunities. For example, almost all districts (each of which is composed of several zones) were found to have a single zone (or small cluster of adjacent zones) that had a large retail employment relative to the others. This concentration of retail activity (which was easily distinguished from corner stores or other predominantly local retail outlets) could be seen as the general attractor that forms the basis of comparison for the destination alternatives. Thus the level of service for each destination alternative was computed to this "shopping zone."

Beyond the allocation of destination alternatives by the trip matrix, however, additional alternatives were given to all households, based on deductive notions of the perception of alternatives. Specifically, "local" alternatives (intrazone and intradistrict) and the central business district (CBD) were allowed as destination alternatives to all households. Because of the household's almost certain familiarity with these alternatives, they were singled out to be included as part of the perceived set of alternatives for all households. In fact, as would be expected by their more favorable levels of service, intrazone and intradistrict travel were observed quite frequently in the sample (almost 40 percent of all shopping trips). The CBD alternative was defined by an aggregate of 6 downtown districts that are all small in area but represent a dense retail area that would be perceived as a single alternative. For households without an automobile, the set of destinations was reduced to those for which bus access was possible.

The final representations of destination alternatives thus were based primarily on a zone system but were adjusted to a more appropriate set corresponding to the more perceptual notion of activity sites by adding specific alternatives that were assumed to be highly attractive alternatives for all households. The inclusion of intrazone and intradistrict alternatives is justified further by the expanding web of perception notion that describes an individual's spatial perceptions as being most detailed in the region immediately surrounding his or her residence and progressively less complete and more aggregate as distance from the home increases.

SPECIFICATION OF UTILITY FUNCTIONS

Table 1 gives the variables and their codes and definitions that were used in the utility functions to describe the 3 choice dimensions. [Annual household income data were coded according to the following classes (in 1968 dollars): (a) 0 to 2,999, (b) 3,000 to 3,999, (c) 4,000 to 5,999, (d) 6,000 to 7,999, (e) 8,000 to 9,999, (f) 10,000 to 11,999, (g) 12,000 to 14,999, (h) 15,000 to 19,999, (i) 20,000 to 24,999, and (j) more than 25,000.] This specification of the joint-choice model for frequency, destination, and mode initially was based in its destination-modal components on a specification developed by Ben-Akiva (2). Several changes were made from that specification, however, to arrive at a form that seemed to better fit the expanded set of choice dimensions being modeled. As a first step, 2 of the level-of-service variables, those representing excess and in-vehicle times, were restructured. The disutility perceived from excess time was seen to be affected by the length of the trip being made. For example, a waiting time of 20 min would seem more onerous on a 1-mile (1.6-km) trip than on a 10-mile (16-km) trip. This was represented by a variable that is formulated as excess time divided by distance for the trip. The other level-of-service variable that was changed for this specification was IVTT, which had been included directly as a measure of disutility. This was restructured, however, as a total travel time (excess plus in-vehicle time) to be included in a logarithmic form in the utility function. Behaviorally, this corresponds to the hypothesis that the sensitivity to absolute changes in total travel time decreases for longer trips.

Figure 1. Creating the estimation sample.

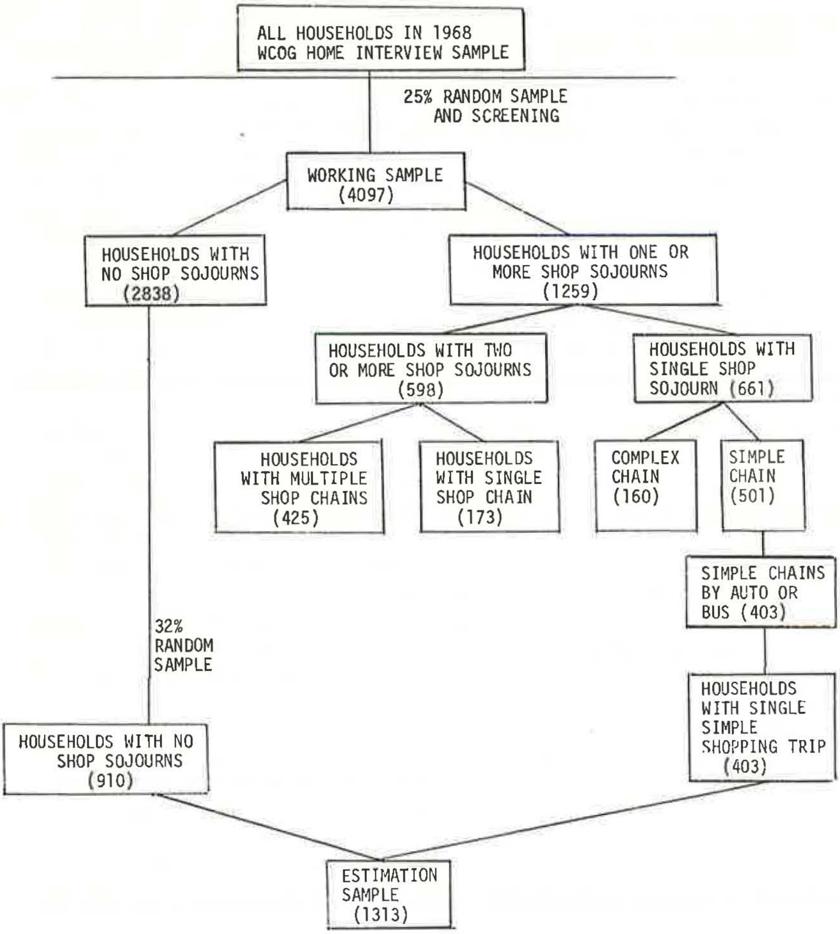


Table 1. Definitions of variables and constants.

Number	Code	Definition
1	DC	1 for car, 0 otherwise
2	OVTT/DIST	Round-trip out-of-vehicle travel time in minutes/1-way distance in miles (kilometers)
3	IVTT + OVTT	Round-trip in-vehicle travel time in minutes + round-trip out-of-vehicle travel time in minutes
4	OPTC/INC	Round-trip out-of-pocket travel cost in cents/annual household income
5	AAC	Number of automobiles available to household - number of automobiles used for work trips by workers in household for car, 0 otherwise
6	1/DIST	1/1-way distance in miles (kilometers)
7	REMP	Retail employment of shopping destination in number of employees
8	DCBD	1 for CBD shopping destination, 0 otherwise
9	DF	1 for 0 frequency, 0 otherwise
10	HHSF	Number of persons in household for 0 frequency, 0 otherwise
11	DENF	Retail employment density in residence zone in employees per acre (hectometer^2) for 0 frequency, 0 otherwise
12	INCF	Annual household income for 0 frequency, 0 otherwise

Note: Alternatives are no trip = 0 frequency; items 1 through 8 = 0; trip to shopping destination d and by mode m is for all relevant shopping destinations including the CBD and for car and transit modes.

Table 2. Utility functions for choice alternatives.

Alternative	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_9	θ_{10}	θ_{11}	θ_{12}
f = 0	0	0	0	0	0	0	0	0	1	HHSF	DENF	INCF
f = 1, m = auto, d = CBD	1	OVTT/DIST	$\ln$(OVTT + IVTT)	OPTC/INC	ACC	1/DIST	$\ln$(REMP)	1	0	0	0	0
f = 1, m = auto, d = nonCBD	1	OVTT/DIST	$\ln$(OVTT + IVTT)	OPTC/INC	ACC	1/DIST	$\ln$(REMP)	0	0	0	0	0
f = 1, m = bus, d = CBD	0	OVTT/DIST	$\ln$(OVTT + IVTT)	OPTC/INC	ACC	1/DIST	$\ln$(REMP)	1	0	0	0	0
f = 1, m = bus, d = nonCBD	0	OVTT/DIST	$\ln$(OVTT + IVTT)	OPTC/INC	ACC	1/DIST	$\ln$(REMP)	0	0	0	0	0

The second major modification to the model form used by Ben-Akiva (2) was the inclusion of a variable representing automobile availability for daytime shopping trips (AAC). This variable is used to explain both frequency and modal choice (more cars available should mean increased likelihood of going shopping and of using a car). This represents one of the areas of complementarity in household decisions that is important to their travel choices: the allocation of automobiles in a household among the variety of household activities. The behavioral hypothesis that leads to inclusion of this form of the variable in the shopping model is that the allocation of automobiles among activities begins with work trips (which are more regular and more important to the household than discretionary purposes) and that the number of automobiles available for daytime shopping is conditional on this choice of mode to work.

A third change from the initial specification was a reformulation of the destination-attraction variables. A logarithmic form of the attraction variable represented by retail employment was chosen because of the large relative value of CBD employment, which caused a high negative value on the CBD dummy variable. Also, instead of using only the retail employment of alternative sites to indicate preferences for larger and more diverse shopping areas over smaller areas, a second destination-specific factor was added to explain destination choice. This was in the form of a variable that is the inverse of 1-way distance from the home zone to the shopping area. The attraction of alternative shopping destinations now is expressed (by the destination-specific factors) as arising from both its relative size and its proximity to the household. This is justified by the hypothesis that a household's knowledge of alternative shopping areas depends on how close it is to them; closer shopping opportunities are more attractive (even beyond the fact that levels of service to them are better) because the household more likely will have better information about the nature of shopping opportunities available and generally will be more likely to actively consider them in its choice set.

One final detail of the specification was that 4 frequency-specific variables were included (along with the level-of-service and attraction variables) to explain frequency choice. These correspond closely to the general classes of variables recommended earlier. They represent household size, ability to stock larger quantities of goods (income), and walking accessibility to shopping alternatives.

The utility functions for the choice alternatives given this specification are given in Table 2. As can be seen, the 3 level-of-service variables enter at positive levels for all alternatives except for 0 frequency when they are also 0. The alternative-specific variables such as automobile availability enter only for the given alternative (in this case, automobile). The dummy variables are 0 or 1 for the specified alternatives.

ESTIMATION RESULTS

Coefficients and other information for this specification of the model are given in Table 3. All of the important policy variables are significant at the 99 percent confidence level. The coefficients of the level-of-service variables (time and cost) have the expected negative signs and result in reasonable values of time. For a typical shopping trip of 2.5 miles (4 km) and a total round-trip travel time of 40 min, OVTT has a disutility that is about twice that of IVTT. Of the attraction factors, 1/DIST has the expected positive sign, indicating that closer destinations are, overall, preferred to those that are more remote. The AAC variable has a relatively large positive parameter, showing that the greater the number of automobiles available is the more likely the household is to make a shopping trip and use the automobile for it.

Of the frequency variables, HHSF has a negative sign, indicating (as expected) that, for a larger household, the probability of not making a shopping trip on a given day becomes less. The variable formulated as DENF is an attempt to account for walk trips to shop, which are not recorded in the home-interview survey. This variable is a proxy for the availability of suitable shopping destinations within walking distance of the home. The expected positive sign of the coefficient of this variable means that a household living in a zone with dense retail employment is more likely not to embark on a vehicle-shopping trip (but is likely to choose a walk-shopping trip instead). The positive sign on the INCF variable, consistent with results obtained elsewhere (4), indicates that higher income households are able to maintain larger stocks of goods and thus will shop with lower frequency.

By using this model specification, we estimated 2 of the important conditional-choice models. Table 5 gives for the conditions in Table 4 the number of observations, number of alternatives, log likelihood for coefficients of 0 $L^*(0)$, log likelihood for estimated coefficients $L^*(\theta)$, and explained log likelihood/total log likelihood ρ^2. In Table 4, parameter estimates for the conditionals of mode given frequency and destination, destination given mode and frequency, and the joint choice of frequency, destination, and mode are compared. As expected and as previously demonstrated by Ben-Akiva (2), the estimated coefficients show great variability depending on the structure used for estimation. The 2 conditionals use less information than the joint estimation uses, and they can be expected to be generally less reliable, theoretically as well as statistically, than the parameters from the joint estimation.

EMPIRICAL APPLICATION OF SHOPPING JOINT-CHOICE MODEL TO EVALUATION OF TRANSPORTATION POLICY OPTIONS

The purpose of this section is to demonstrate how the disaggregate choice models developed in this research can be used to illustrate the behavioral effects of various transportation options on the demand for travel. The emphasis in this analysis was on highlighting the varying effects that several transportation alternatives would have on 3 different types of households. The types of households were represented by 3 of the samples taken in the 1968 Washington, D.C., home-interview survey. A similar analysis could be performed by using a random sampling of all households in the area or by constructing segments on which the effects could be compared. A larger scale case study, which also is being conducted in this research project, is using a set of disaggregate models, including this one, in a full network equilibration framework. The use of 3 typical households (for which frequency, mode, and destination, but not route choice, were forecast) was chosen for this study primarily for reasons of simplicity of presentation and ease of computation.

Policy Alternatives

Five policy alternatives were chosen to be compared to the base case. These alternatives are representative of the range of options currently being considered in response to, among other issues, air-quality and energy-conservation programs. The alternatives can be summarized as follows:

1. Base case—conditions existing in Washington, D.C., in 1968;
2. Case 1—gasoline prices 3 times greater than those of 1968;
3. Case 2—parking costs 3 times greater than those of 1968;
4. Case 3—employer-based car-pool incentives and special car-pool lanes to decrease travel time to work to 70 percent of 1968 base times;
5. Case 4—transit available for all trips [IVTT as good as for automobile and OVTT = 20 min + (10 min × number of transfers) but no more than existing conditions]; and
6. Case 5—combination of cases 1 through 4.

Typical Households

Three households were chosen from the Washington, D.C., area to represent a range of characteristics, from low-income, captive-transit inner-city residents to high-income, automobile-captive suburban residents. The specific characteristics of these households that are relevant as inputs to the model are given in Table 6. Not given (but used

Table 3. Model coefficient values.

Variable	Coefficient	t-Statistic	Variable	Coefficient	t-Statistic
DC	-0.555	-2.13	ℓn(REMP)	0.161	3.29
OVTT/DIST	-0.100	-3.38	DCBD	0.562	2.07
ℓn(IVTT + OVTT)	-2.24	-11.85	DF	-3.78	-4.51
OPTC/INC	-0.0242	-4.20	HHSF, 0 frequency only	-0.186	-4.57
AAC, car only	0.557	5.61	DENF, 0 frequency only	0.383	1.38
1/DIST	0.0686	1.66	INCF, 0 frequency only	0.0414	1.18

Note: Number of observations = 1,313; number of alternatives = 44,718; log likelihood for coefficients of 0 = -3,830; log likelihood for estimated coefficients = -2,511; and explained log likelihood/total log likelihood = 0.36.

Table 4. Comparison of conditionals and joint estimation.

Variable or Constant	(m|f, d) Estimate	(m|f, d) Standard Error	(d|f, m) Estimate	(d|f, m) Standard Error	(f, d, m) Estimate	(f, d, m) Standard Error
DC	-1.35	0.732			-0.555	0.260
OVTT/DIST	-0.116	0.623	-0.0399	0.0277	-0.100	0.0296
ℓn(OVTT + IVTT)	-2.21	0.367	-2.60	0.240	-2.24	0.189
OPTC/INC	-0.0243	0.0151	-0.0237	0.00726	-0.024	0.00576
AAC	1.63	0.667			0.557	0.0992
1/DIST			0.0344	0.0634	0.0686	0.0414
ℓn(REMP)			0.370	0.0533	0.161	0.0489
DCBD			0.354	0.284	0.562	0.271
DF					-3.78	0.839
HHSF					-0.186	0.0404
DENF					0.383	0.276
INCF					0.0414	0.0350

Table 5. Log functions and other data for conditionals and joint estimation of Table 4.

Condition	Number of Observations	Number of Alternatives	L*(0)	L*(θ)	ρ^2
(m|f, d)	225	450	-156	-62	0.60
(d|f, m)	403	8,732	-1,210	-988	0.19
(f, d, m)	1,313	44,718	-3,830	-2,511	0.36

Table 6. Typical households.

Characteristic	Household 1	Household 2	Household 3
Income per year, dollars	5,000	13,500	22,500
Number of automobiles owned	0	1	2
Household size	4	4	5
Distance to CBD, miles	3	6	9.5
Retail employment density at residential zone	High	Medium	Low
Transit availability	Good	Medium	None
Distance to work, miles	2.5	4	6.5

Note: 1 mile = 1.6 km.

Table 7. Model forecasts for households 1, 2, and 3.

Household	Alternative	Mode to Work (percentage of individuals) Drive Alone	Car Pool	Transit	Mode to Shop (household trips/day) Car	Transit	Total
1	Base case	0	13.8	86.2	0	0.010	0.010
	Case 1	0	12.0	88.0	0	0.010	0.010
	Case 2	0	13.8	86.2	0	0.010	0.010
	Case 3	0	42.2	57.8	0	0.010	0.010
	Case 4	0	7.6	92.4	0	0.072	0.072
	Case 5	0	37.6	62.4	0	0.072	0.072
2	Base case	89.4	2.9	7.7	0.482	0.012	0.494
	Case 1	89.1	1.4	9.5	0.420	0.013	0.433
	Case 2	89.4	2.9	7.7	0.454	0.013	0.467
	Case 3	81.2	11.9	6.9	0.494	0.012	0.506
	Case 4	76.7	2.6	20.7	0.463	0.083	0.546
	Case 5	65.8	11.9	22.3	0.389	0.094	0.483
3	Base case	96.2	3.8	0	0.499	0	0.499
	Case 1	95.7	4.3	0	0.441	0	0.441
	Case 2	95.4	4.6	0	0.478	0	0.478
	Case 3	83.2	16.8	0	0.517	0	0.517
	Case 4	82.4	3.3	14.7	0.468	0.096	0.564
	Case 5	63.7	17.7	18.6	0.418	0.105	0.523

in the model calculations) are the residence locations of the households, which affect the set of relevant shopping destination and modal alternatives. Also not given are the work locations, which, however, are notable only in that household 3 alone has a worker who commutes to a downtown location that requires a parking fee.

Use of Models to Forecast Effects of Transportation Options

The 2 models that are used for this application are the joint model of choice of mode to shop previously described and model of choice of mode to work (also multinomial logit) developed in related research that forecasts the probability of choice among the automobile-driven-alone, car-pool, and bus modes (5). The choice-of-mode-to-work model takes as inputs level-of-service and modal-specific variables similar to those used in the shopping model. To distinguish the effects of car-pool incentives on the use of car-pools, we included a variable indicating the presence of employer-based car-pool incentives (as they existed in 1968 for government workers) in the model.

The forecasting procedure that was used for each policy alternative is as follows: First, the independent variables were introduced in the choice-of-mode-to-work model to produce forecasts of modal-choice probabilities; then, the shopping model was applied with the independent variables including the residual automobile-availability variable that resulted from the forecast probabilities of choice of mode to work. This procedure was repeated for each household.

The forecasts of the models for all the policy alternatives for each household are given in Table 7. The 3 modal probabilities for mode to work reflect the availability as well as the characteristics of the alternatives: Car pool was allowed for everyone, drive alone was allowed only for those owning automobiles, and transit was allowed only where it was available. The number of household shopping trips per day is given directly by the model, and this is multiplied by the modal probabilities (also taken directly from the model output) to give trips per day by automobile and bus (car pool is not a relevant mode for shopping).

Results of Comparisons

Tables 8, 9, and 10 give a summary of the important results of the analysis. The data given in Table 8 show how the number of shopping trips made by each household varies for the 6 alternatives (trip frequency for the work trip, of course, remains constant). The first household, which is captive to transit, makes more shopping trips only when transit improvements are implemented (increases more than 6 times over base levels). For the second household, decreases are observed, as expected with price disincentives on automobile use. However, the introduction of car-pool incentives shifts use of the household's automobile away from the work trip and leaves it available for daytime shopping by other household members. Thus, car-pool incentives increase the amount of shopping travel by automobile. Transit improvements increase shopping travel by transit, as expected, for all households. The data given in Table 9 translate the shopping and work travel into daily vehicle miles of travel (VMT) (vehicle kilometers of travel) for each household (based on forecast probabilities and known distances to alternative destinations). As expected, the 2 price increases on automobile travel reduce VMT (vehicle kilometers of travel) for all households for both work and shopping travel. Car-pool incentives, however, increase VMT (vehicle kilometers of travel) for shopping in those households that own automobiles and increase work-trip VMT (vehicle kilometers of travel) for households that previously had no direct access to an automobile for travel to work, but now are encouraged to use car pools. The total VMT (vehicle kilometers of travel) over all 3 households and across the 2 trip purposes of work and shopping is less than for the base case; however, inclusion of other travel purposes, such as recreation and personal business, which also are affected by automobile availability, easily could make the total VMT (vehicle kilometers of travel) for the car-pool-incentives option greater than that for the base case.

Table 8. Effect of policy alternatives on household shopping trips.

	Household 1		Household 2		Household 3	
Alternative	Household Shopping Trips/Day	Change From Base (percent)	Household Shopping Trips/Day	Change From Base (percent)	Household Shopping Trips/Day	Change From Base (percent)
Base case	0.010	0	0.494	0	0.499	0
Case 1	0.010	0	0.433	-12	0.441	-12
Case 2	0.010	0	0.467	-5	0.478	-4
Case 3	0.010	0	0.506	+3	0.517	+4
Case 4	0.072	+620	0.546	+10	0.564	+13
Case 5	0.072	+620	0.483	-2	0.523	+5

Table 9. Effect of policy alternatives on vehicle miles (kilometers) of travel.

		Work		Shop		Total	
Household	Alternative	VMT	Change From Base (percent)	VMT	Change From Base (percent)	VMT	Change From Base (percent)
1	Base case	0.276	0	0	0	0.276	0
	Case 1	0.240	-13	0	0	0.240	0
	Case 2	0.276	0	0	0	0.276	0
	Case 3	0.844	+206	0	0	0.844	+206
	Case 4	0.152	-45	0	0	0.152	-45
	Case 5	0.752	+172	0	0	0.752	+172
2	Base case	7.245	0	4.020	0	11.265	0
	Case 1	7.173	-1	3.327	-17	10.500	-7
	Case 2	7.245	0	3.844	-4	11.089	-2
	Case 3	6.877	-5	4.127	+3	11.004	-2
	Case 4	6.219	-14	3.870	-4	10.089	-10
	Case 5	5.645	-22	3.127	-22	8.772	-22
3	Base case	12.70	0	4.48	0	17.18	0
	Case 1	12.66	-0.3	3.62	-19	16.28	-5
	Case 2	12.64	-0.5	4.26	-5	16.90	-2
	Case 3	11.69	-8	4.64	+4	16.33	-5
	Case 4	10.88	-14	4.21	-6	15.09	-12
	Case 5	9.20	-28	3.41	-24	12.61	-27
1, 2, and 3	Base case	20.22	0	8.50	0	28.72	0
	Case 1	20.07	-1	6.95	-18	27.02	-6
	Case 2	20.16	0	8.10	-5	28.26	-2
	Case 3	19.41	-4	8.77	+3	28.18	-2
	Case 4	17.25	-15	8.08	-5	25.33	-12
	Case 5	15.60	-23	6.54	-23	22.14	-23

Note: 1 vehicle mile of travel = 1.6 vehicle km of travel.

Table 10. Bias from assuming no effect of level-of-service changes on frequency of household shopping trips.

	Household 1			Household 2			Household 3		
Alternative	Actual VMT	VMT Given Base-Trip Generation	Bias (percent)	Actual VMT	VMT Given Base-Trip Generation	Bias (percent)	Actual VMT	VMT Given Base-Trip Generation	Bias (percent)
Base case	0			4.020			4.480		
Case 1	0	0	0	3.327	3.729	+12	3.623	4.095	+13
Case 2	0	0	0	3.844	4.065	+6	4.263	4.445	+4
Case 3	0	0	0	4.127	4.025	-2	4.642	4.480	-3
Case 4	0	0	0	3.870	3.500	-10	4.205	3.716	-12
Case 5	0	0	0	3.127	3.197	+2	3.407	3.249	-5

Note: 1 vehicle mile of travel = 1.6 vehicle km of travel.

The data given in Table 10 show the bias that results from assuming that level-of-service changes have no effect on the frequency with which households make shopping trips. Actual VMT (vehicle kilometers of travel) given base-trip data are as predicted by the full joint-choice model; the VMT (vehicle kilometers of travel) given base-trip data are as computed by the conditional model $P(d, m \mid f_{base})$, which assumes frequency to be unaffected by the transportation options. The bias percentages are those that result in this application of the model. Although the biases from not including price effects on travel frequency are not overwhelming in magnitude, they are consistent in underestimating the effectiveness of disincentives to decrease automobile use and of transit improvements to increase public transit patronage. Thus, to accurately model these policy alternatives, which often have impacts on demand only on the order of magnitude of the observed biases, it would seem extremely important to consider, in the model structure, the effects of level-of-service changes on frequency of travel.

CONCLUSIONS AND IMPLICATIONS OF RESULTS

The estimation of this joint-choice model provides what is an encouraging, though not final, step in the development of a full set of disaggregate choice models of travel demand. The relative ease of estimating the single joint-choice model compared to calibrating 3 separate models by using arbitrary sequence assumptions is a clear advantage (even beyond the general acceptability of joint over conditional models). All the estimated coefficients have reasonable signs and magnitudes and relatively small standard errors that primarily are due to the full use of the data by joint estimation. The joint model was estimated by using a maximum likelihood procedure that consumed less than 1 min of central processing unit time in 80,000 bytes of core on an IBM 370/165.

Several important properties of the disaggregate choice model set were demonstrated in the example application of the models. It was shown how the models can be used directly to compute the quantitative effects of transportation policy options on the travel demand of either specific types of households or of more generally constructed market segments. The inclusion of a large set of policy-relevant variables in the model specification allows for the testing of a wide range of options, and the model forecasts can be computed directly for many types of impacts—from the effect on CBD shopping frequency of parking cost increases to the effect of car-pool incentives on areawide VMT (vehicle kilometers of travel).

Another set of properties that were demonstrated in the case study were some of the effects of model structure on the resulting forecasts. One feature of the model set used here is the explicit linking of household decisions in choosing among travel alternatives. The choice situation that is represented in these models is the automobile-allocation decision: whether the automobile will be used for the work trip and how this affects household choices for other types of trips. That an automobile left at home will stimulate automobile travel for discretionary purposes (shopping) is an extremely important effect in evaluating car-pool incentive programs. Other household travel choices that involve complementarity among trip purposes (such as the consolidation of travel through trip chaining) can and should be similarly represented to present a more complete behavioral picture of travel demand.

Another structural property of the shopping joint-choice model that has been shown to be important is the representation of level-of-service effects on travel frequency. The bias from not including this effect is significant in that it is consistent in underestimating the effectiveness of some of the currently relevant transportation options. Thus use of a model structure that represents effects of levels of service on all choices (in this case, a joint-choice structure) is necessary to realistically appraise policy options.

Two areas of further work are logical continuations of the effort documented here. The first is in the development of improved specifications to increase the sensitivity of the model toward a larger variety of policy options. This might include extension of the model to treatment of additional modes or simply inclusion of additional variables to either strengthen the behavioral representation or allow an expanded set of policy variables. A group of alternative specifications for which estimation already has been performed is documented by Adler (1).

A more fundamental gap between current travel demand models and existing theories of travel behavior still exists. All modeling efforts to date have stratified trips by trip purpose and used a single link as the unit of travel demand. The most recent choice models such as the one reported in this paper have assumed simple round trips as the relevant unit. It is becoming increasingly obvious, however, that estimation of these models is not adequate to describe the large numbers of more complex trip patterns that are observed in urban travel. In the 1968 Washington, D.C., home-interview survey, patterns of travel that integrated shopping trips with other trip purposes (shopping on the way home from work) are observed in greater numbers than the simple patterns in which a household makes a single 2-link round trip for shopping. The recent trend, which largely is due to rising fuel prices, has been toward increased traveler tendencies to consolidate their needs for transportation by linking several purposes in a single, expanded round trip from home. There are several trade-offs involved in the households' comparisons among travel patterns. One is the desire to satisfy needs for travel as they accumulate to a threshold level against the attempt to unite them temporally to allow for a single, more efficient round trip. Clearly, there are behavioral issues here that have potentially great impact on energy-conservation programs as well as on other modally oriented incentives (relative advantages of the various modes in servicing these more complex patterns of travel) but remain unaddressed in any current travel demand models.

The problem with the most recent efforts in addressing the issues posed by complex patterns of travel seems to be their orientation around single trip purposes as a means of stratification of behavioral responses. A more integrated approach would be in the use, as a unit of demand, of complete patterns of household travel and in the identification of the general classes of travel that can be decomposed from those patterns. For example, a useful classification might distinguish among fixed patterns of travel (such as travel to work or school where the destination is generally static in the short term), discretionary travel (where mode, destination, and frequency are active choices), and patterns where fixed and discretionary travel purposes are combined (as in a shopping trip on the way home from work). Such a scheme would allow for behavioral comparisons among all patterns of household travel rather than exclude the more complex patterns that are of increasingly greater interest to transportation planners as the more restrictive travel-purpose-based stratifications do.

This expansion of the scope of disaggregate behavioral models will, of course, benefit from the research of the past few years. In particular, the general format of the joint-choice model is seen as being a key to the modeling of dimensions of choice (for example, among morphologically different multipurpose round trips) that are even less subject to the imposition of a sequence assumption than the frequency, mode, and destination choices now being modeled.

ACKNOWLEDGMENTS

The research described in this paper was performed at the Massachusetts Institute of Technology under a contract with the University Grant Research Program of the U.S. Department of Transportation. The title of the full research is Experiments to Clarify Priorities in Urban Travel Forecasting Research and Development. The views and opinions expressed in this paper are ours and do not necessarily represent those of the U.S. Department of Transportation.

The assistance of the Department of Transportation Planning of the Metropolitan Washington (D.C.) Council of Governments in providing the necessary data for the project is appreciated.

REFERENCES

1. T. J. Adler. A Joint Choice Disaggregate Model of Non-Work Urban Passenger Travel Demand. Department of Civil Engineering, Massachusetts Institute of Technology, MS thesis.
2. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, Massachusetts Institute of Technology, PhD dissertation, 1973.
3. M. E. Ben-Akiva. Structure of Passenger Travel Demand Models. Transportation Research Record 526, 1974, pp. 26-42.

4. A Disaggregate Behavioral Model of Urban Travel Demand. Charles River Associates, Inc., 1972.
5. S. W. Haws and M. Ben-Akiva. Estimation of a Work Mode Choice Model Which Includes the Carpool Mode. Department of Civil Engineering, Massachusetts Institute of Technology, Working Paper DOT-WP-4, 1974.
6. M. G. Richards and M. E. Ben-Akiva. A Disaggregate Travel Demand Model. Saxon House, D. C. Heath, Ltd., England, 1975.
7. O. Westelius. The Individual's Pattern of Travel in an Urban Area. National Swedish Institute for Building Research, Stockholm, 1972.

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Charles V. Wootan, Texas A&M University, chairman

TRANSPORTATION FORECASTING SECTION
E. Wilson Campbell, New York State Department of Transportation, chairman

Committee on Passenger Travel Demand Forecasting
Daniel Brand, Harvard University, chairman
Moshe Ben-Akiva, Robert B. Dial, Raymond H. Ellis, David S. Gendell, Thomas F. Golob, John R. Hamburg, Lawrence V. Hammel, Walter G. Hansen, David T. Hartgen, Kevin E. Heanue, Frank E. Horton, Louis E. Keefer, C. E. Klamm, Gerald Kraft, Eugene J. Lessieu, Herbert S. Levinson, William McGrath, James M. McLynn, Sydney R. Robertson, Peter R. Stopher, Vergil G. Stover, Darwin G. Stuart, Anthony R. Tomazinis, Edward Weiner, Charles C. Whitmarsh, George V. Wickstrom, Alan G. Wilson

James A. Scott, Transportation Research Board staff

The organizational units and the chairmen and members are as of December 31, 1974.