TRANSPORTATION RESEARCH RECORD
526

Formerly issued as Highway Research Record

Travel Demand Forecasting

7 reports prepared for the 53rd Annual Meeting of the Highway Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

Washington, D. C., 1974

Transportation Research Record 526
Price $4.40
Edited for TRB by Marjorie Moore

subject areas
55 traffic measurements
84 urban transportation systems

Transportation Research Board publications are available by ordering directly from the Board. They are also obtainable 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.

These papers report research work of the authors that was done at institutions named by the authors. The papers were offered to the Transportation Research Board of the National Research Council for publication and are published here in the interest of the dissemination of information from research, one of the major functions of the Transportation Research Board.

Before publication, each paper was reviewed by members of the TRB committee named as its sponsor and accepted as objective, useful, and suitable for publication by the National Research Council. The members of the review committee were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the subject concerned.

Responsibility for the publication of these reports rests with the sponsoring committee. However, the opinions and conclusions expressed in the reports are those of the individual authors and not necessarily those of the sponsoring committee, the Transportation Research Board, or the National Research Council.

Each report is reviewed and processed according to the procedures established and monitored by the Report Review Committee of the National Academy of Sciences. Distribution of the report is approved by the President of the Academy upon satisfactory completion of the review process.

The National Research Council is the principal operating agency of the National Academy of Sciences and the National Academy of Engineering, serving government and other organizations. The Transportation Research Board evolved from the 54-year-old Highway Research Board. The TRB incorporates all former HRB activities but also performs additional functions under a broader scope involving all modes of transportation and the interactions of transportation with society.

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA

National Research Council. Highway Research Board.
Travel demand forecasting.

(Transportation research record; 526)

1. Traffic surveys—Congresses. 2. Urban transportation—Congresses. I. National Research Council. Transportation Research Board. II. Title. III. Series.
TE7.H5 no. 526 [HE370] 380.5'08s [388.4'131'4] 75-6802
ISBN 0-309-02373-4

CONTENTS

FOREWORD iv

ESTIMATING THE DEMAND FOR SHORT-HAUL AIR TRANSPORT SYSTEMS
Adib K. Kanafani and Shing-Leung Fan 1

DETERMINATION OF FUNCTIONAL SUBREGIONS WITHIN AN URBAN AREA FOR TRANSPORTATION PLANNING
Thomas F. Golob, Salvatore J. Hepper, and John J. Pershing, Jr. 16

ALTERNATIVE TRAVEL BEHAVIOR STRUCTURES
Structure of Passenger Travel Demand Models
Moshe E. Ben-Akiva 26
Disaggregate Access Mode and Station Choice Models for Rail Trips
Peter S. Liou and Antti P. Talvitie 42
Discussion
Gerald Kraft 52
Peter R. Stopher 56
David T. Hartgen 57
Authors' Closure
Moshe Ben-Akiva 59
Peter S. Liou and Antti P. Talvitie 64

THE NATURE OF TRAVEL DECISION-MAKING
Reid H. Ewing 66

STRUCTURAL MODEL FOR EVALUATING URBAN TRAVEL RELATIONSHIPS
Edward J. Kannel and Kenneth W. Heathington 73

A COMPARATIVE EVALUATION OF INTERCITY MODAL-SPLIT MODELS
John C. Bennett, Raymond H. Ellis, John C. Prokopy, and Melvyn D. Cheslow 83

SYNTHETIC DERIVATION OF INTERNAL TRIPS FOR SMALL CITIES
John W. Bates 93

SPONSORSHIP OF THIS RECORD 105

FOREWORD

This RECORD will be of great interest to researchers and practitioners of travel demand forecasting because it presents findings from leading researchers in the field on recently recognized interactions between travel behavior assumptions and the resulting travel models and forecasts made with these assumptions.

Travel forecasters have always recognized that their forecasts should be plausible and reasonable and that their models must be calibrated to be consistent with observed travel patterns. What emerges from the findings of the 1972 Conference on Urban Travel Demand Forecasting (HRB Special Report 143) and now from the papers in this RECORD is that the forecasting models being used in the profession for the last 20 years have their own implicit rules of behavior. These behavior rules are not only those implied by the independent variables included and the coefficients estimated but also those on the structure of the travel choices (e.g., sequence of choices) and choice alternatives (e.g., alternative destinations) over which the models should be applied. It is the latter rules of behavior required by the current travel forecasting models (e.g., trip generation equations, gravity and logit models) and our method of estimating these models that require exposure to the profession. A full understanding of the choice structure implied by the model being used can be of great help in efficient use of the model and in appropriate selection of what forecasting model to use in the first place. Lack of knowledge of a model's implied choice behavior may lead to grievous errors in its application. Lack of knowledge of how travelers behave under different circumstances when confronted with high-capital versus low-capital transportation alternatives may require that complicated models always be used.

Clearly, more research on travel behavior and behavioral travel models is required before we can make trade-offs 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. Examination of the papers in this RECORD will carry us a substantial distance forward in understanding that these trade-offs can be made and in beginning to make these trade-offs in our own work.

—Daniel Brand

ESTIMATING THE DEMAND FOR SHORT-HAUL AIR TRANSPORT SYSTEMS

Adib K. Kanafani and Shing-Leung Fan,
Institute of Transportation and Traffic Engineering,
University of California, Berkeley

To evaluate the feasibility of novel systems for short-haul air transportation requires an estimation of the market share potential for various configurations of such systems. This paper deals with the development of a model for estimating the market share that various short take-off and landing (STOL) system configurations can be expected to capture in a high-density, short-haul air travel corridor. The process by which travelers in the corridor choose among different routes serving the corridor is studied. Variables such as line-haul travel times, schedule frequencies, and fares are studied. Traveler's choice is modeled in terms of these variables in a probabilistic manner. Such a formulation allows the aggregation of travelers into groups for the purpose of demand analysis. The model is calibrated on the basis of data on travel characteristics in the 500-mile corridor connecting the San Francisco Bay and Los Angeles metropolitan areas. System configurations include STOLports located at various points within the region and varying schedule frequencies and air fares. Alternative strategies of diverting short-haul air traffic from congested hub airports to STOLports are also studied. The calibrated choice model is combined with a total travel forecasting model to provide a forecasting procedure for estimating the demand potential for STOL transportation systems. The calibrated models are used to study various STOL system configurations and to estimate their market potential.

•THE objective of the research documented in this paper was to develop a procedure for forecasting the demand for alternative short take-off and landing (STOL) systems in the high-density, short-haul air travel corridor connecting San Francisco and Los Angeles.

The forecasting framework used has two stages. In the first stage total air travel demand in the corridor is forecast. This is followed by the second stage of estimating the choice among available air travel routes in the corridor. The combination of the two allows the estimation of the demand for any one route or type of service, including a variety of postulated STOL systems. This distinction between total air travel demand and the choice among available routes reflects a characteristic particular to short-haul air transportation. In short-haul air transportation (normally defined by a range of approximately 500 miles), line-haul travel time does not constitute the major portion of total travel time. The total quality of air service is, therefore, more sensitive to variations in ground access travel times and schedule delays than is the case in long-haul transport. Consequently, it is necessary to study the process of choice among alternative routes and to relate that process to such route characteristics as access times and schedule frequency.

Another characteristic peculiar to short-haul air transportation is that within its range high-speed ground transportation modes may pose significant competition. In principle, therefore, a forecasting procedure for short-haul air transportation should consider the interplay among all available air and ground modes. However, for this study of the California corridor, it was believed that ground transportation technology has not yet reached a point where significant interaction occurs between air and ground transportation. Consequently, the forecasting process developed deals exclusively with air transportation. It is therefore clear that the forecasts obtained here are conditional on the fact that no significant changes occur in the ground transportation system in the corridor. Should a high-speed rail system, for example, become a reality for the California corridor, then the forecasts presented here would be false.

MODEL DESIGN

The modeling structure used in this study is shown in Figure 1. A number of travel routes are identified within the short-haul corridor, and for each route a number of attributes or transportation characteristics are identified. Two models are used to estimate the demand potential for STOL transport in the corridor. First, a generation model is used to estimate the total air travel demand, where it is postulated that this demand depends on the socioeconomic characteristics of the city pairs in the corridor as well as the best route attributes available for each city pair. Second, a choice model is used to estimate the split in the total demand among available routes in the corridor, where it is assumed that the split depends on the relative attributes of each of the routes. These two models are then combined to estimate the total market share for each route. When summed over routes that constitute STOL service, the total demand for STOL transportation is obtained.

The study corridor on which these models were to be applied was represented by a network consisting of origin and destination cities and origin and destination airports. As shown in Figure 2, the corridor joins two regions—I, San Francisco, and II, Los Angeles—with air transport among a number of airport pairs. For every origin-destination pair, a route is defined by a path along the network extending from the origin city to the origin airport, the destination airport, and finally the destination city. In Figure 2 ACDB and AEKB are examples of routes connecting cities A and B.

Figure 1. Model framework.

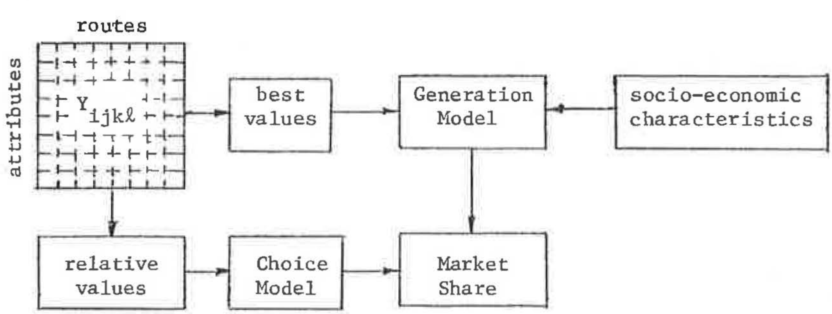
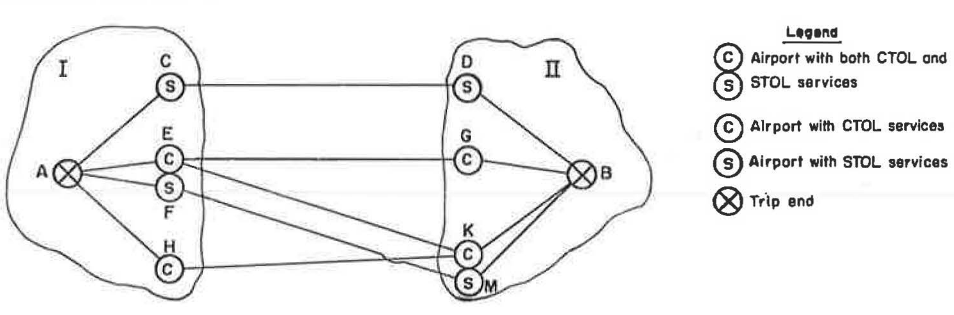


Figure 2. Graphical representation of corridor.



Air Travel Generation Model

A simple multiplicative model was used to estimate intercity air transport demand in the study corridor. This model included both socioeconomic (population, median income, and employment) and transportation (the best available schedule frequency, lowest travel time, and lowest travel cost) variables among all the available routes for each city pair. Three alternative model forms were specified and later statistically tested:

$$T_{ij} = \alpha_0 P_i^{a1} P_j^{a2} Y_i^{a3} Y_j^{a4} t_{ij}^{a5} \quad (1)$$

$$T_{ij} = \alpha_0 P_i^{a1} P_j^{a2} Y_{ij}^{a3} t_{ij}^{a4} \quad (2)$$

$$T_{ij} = \alpha_0 P_i^{a1} P_j^{a2} Y_{ij}^{a3} LS_{ij}^{a4} \quad (3)$$

where

T_{ij} = total traffic between cities i and j,
P = population,
Y = median income,
Y_{ij} = average median income for both cities,
t_{ij} = shortest travel time among all routes between i and j,
LS_{ij} = level-of-service variable defined as the average travel time between i and j over all routes and weighted by the cost and the schedule frequency for each route, and
α = parameter representing demand elasticities with respect to the variables.

Route Choice Model

The purpose of the choice model is to describe how a traveler in the corridor is likely to choose among the available routes serving the corridor. This description is then used to split the total travel demand generated by the previous model among these routes. The model specified in the study was a stochastic model that predicts the probability of choice conditional on values of the choice elasticities. By studying the random variations of these elasticities among individuals and using a procedure proposed by Kanafani (3), we can aggregate over the total study population.

Using the corridor notation of Figure 2, let P_{ijk} be the probability that a traveler between cities i and j chooses route k and let Y_{ijkl}, $l = 1, \ldots, m$, be m attributes of route k. The basic postulate of the model is specified by the following probability function:

$$P_{ijk} = g(Y_{ijk1}, Y_{ijk2}, Y_{ijk3}, \ldots, Y_{ijkm})$$

That is, the probability of choice is assigned on the basis of a set of m route attributes.

An individual is assumed to evaluate the characteristics of all routes one at a time. For each characteristic he ranks the routes available to him. This ranking is analogous to the probability that a route is chosen on the basis of this particular characteristic. Thus it is assumed that there is a unique correspondence between the ranking of a route on the basis of a characteristic and the probability of choosing the route on that basis. This correspondence is defined by a set of weights θ_l. Letting A_l be the event of choosing a route on the basis of characteristic l and postulating a sigmoidal relationship among the weight θ, the value Y_l of l, and the probability P give the choice probability P_{ijkl} as

$$P_{ijkl} = P[A_l] = \frac{Y_{ijkl}^{\theta_l}}{\sum_r Y_{ijrl}^{\theta_l}} \quad (4)$$

In Eq. 4 the probability of taking route k on the basis of attribute l is a function of its value for route k relative to all available routes. It should be noted that P_{ijkl} in Eq. 4 is the probability based only on one attribute, l, and is independent of all the other attributes. That is, $P_{ijk1}, \ldots, P_{ijkm}$ are probabilities of independent events. The total choice probability P_{ijk}, which is based on all route attributes, is therefore

$$P_{ijk} = \prod_{l=1}^{m} P[A_l] = \prod_{l=1}^{m} \left[Y_{ijkl}^{\theta_l} \Big/ \sum_{Y} Y_{ijrl}^{\theta_l} \right] \tag{5}$$

subject to

$$0 \leq P_{ijk} \leq 1 \tag{6}$$

and

$$\sum_{k} P_{ijk} = 1 \tag{7}$$

Equation 7 is satisfied by introducing a factor K_{ij} in Eq. 5 to give

$$P_{ijk} = K_{ij} \prod_{l=1}^{m} \left[Y_{ijkl}^{\theta_l} \Big/ \sum_{Y} Y_{ijrl}^{\theta_l} \right] \tag{8}$$

With Eqs. 7 and 8 it should be possible to determine K_{ij}.

To facilitate the presentation of the remainder of the model, we assume without loss of generality that there are only three route attributes: total travel time H_{ijk}, schedule frequency F_{ijk}, and travel cost C_{ijk}. Equation 8 now becomes

$$P_{ijk} = K_{ij} \left[F_{ijk}^{\alpha} \Big/ \sum_{r} F_{ijr}^{\alpha} \right] \left[C_{ijk}^{\beta} \Big/ \sum_{r} C_{ijr}^{\beta} \right] \left[H_{ijk}^{\gamma} \Big/ \sum_{r} H_{ijr}^{\gamma} \right] \tag{9}$$

where α, β, and γ are the weights placed on each of the attributes. Combining Eqs. 7 and 9 gives

$$K_{ij} = \frac{\sum_{k} F_{ijk}^{\alpha} \sum_{k} C_{ijk}^{\beta} \sum_{k} H_{ijk}^{\gamma}}{\sum_{r} F_{ijr}^{\alpha}\, C_{ijr}^{\beta}\, H_{ijr}^{\gamma}} \tag{10}$$

Substituting this value in Eq. 9 gives the expression of the choice probability, which, because α, β, and γ are postulated as random variables, is stated as a conditional probability of choice given α, β, and γ:

$$P[ijk \mid \alpha, \beta, \gamma] = F_{ijk}^{\alpha}\, C_{ijk}^{\beta}\, H_{ijk}^{\gamma} \Big/ \sum_{r} F_{ijr}^{\alpha}\, C_{ijr}^{\beta}\, H_{ijr}^{\gamma} \tag{11}$$

To find the unconditional probability requires that this expression be integrated over the domains of the random variables α, β, and γ respectively, which gives

$$P[ijk] = \int_{R_1} \int_{R_2} \int_{R_3} P[ijk \mid \alpha, \beta, \gamma] f(\alpha, \beta, \gamma)\, d\alpha d\beta d\gamma \tag{12}$$

where $f(\alpha, \beta, \gamma)$ is the joint density function of the variables α, β, and γ. It was assumed and later statistically verified that a traveler assigns these weights independently of one another. This assumption yields a considerable simplification because it allows the representation of the joint density function as the product of the individual density functions for each weight. The choice model can now be specified in its complete form:

$$P[ijk] = \int_{R_1} \int_{R_2} \int_{R_3} \frac{F_{ijk}^{\alpha}\, C_{ijk}^{\beta}\, H_{ijk}^{\gamma}}{\sum_r F_{ijr}^{\alpha}\, C_{ijr}^{\beta}\, H_{ijr}^{\gamma}} f_1(\alpha) f_2(\beta) f_3(\gamma) d\alpha d\beta d\gamma \tag{13}$$

STOL Share Model

Once both the travel generation model and the choice model are completely specified, we combine them into a model that will allow the estimation of the share of any route in a corridor. This will also allow the estimation of the demand potential for STOL transport. Combining the value T_{ij} of the demand for air travel between any O-D pair, as obtained from the generation model, with the choice probability $P[ijk]$, as obtained from the choice model, gives the expected demand for a route k:

$$E[T_{ijk}] = T_{ij}\, P[ijk] \tag{14}$$

If ψ denotes the subset of all routes k that are STOL routes, then the total demand potential for STOL transportation between any O-D pair i, j can be obtained from

$$E[ST_{ij}] = T_{ij} \sum_{k \in \psi} P[ijk] \tag{15}$$

and the total STOL demand potential in the corridor is obtained by adding the demand values for all O-D pairs:

$$E[ST] = \sum_i \sum_j E[ST_{ij}] \tag{16}$$

This model allows the estimation of the demand potential for STOL transportation for any STOL service configuration.

THE DATA BASE

Most of the data used in this study were derived from an on-board origin-destination survey conducted in 1970 by Daniel, Mann, Johnson, and Mendenhall (1). For each trip the following variables were observed:

1. Trip origin and destination,
2. Airport pair used,
3. Trip purpose, and
4. Reported ground access times at both trip ends.

A total of 1,637 business trips and 1,467 nonbusiness trips were included in the data base. This trip information was collected on 12 conventional take-off and landing (CTOL) routes in the study corridor:

1. Oakland-Hollywood/Burbank,
2. Oakland-Los Angeles International,
3. Oakland-Ontario,
4. Oakland-Santa Ana (Orange County),
5. San Francisco International-Hollywood/Burbank,
6. San Francisco International-Long Beach,
7. San Francisco International-Ontario,
8. San Francisco International-Santa Ana,

9. San Jose-Hollywood/Burbank,
10. San Jose-Los Angeles International,
11. San Jose-Ontario, and
12. San Jose-Santa Ana.

A major deficiency of the data source was that the survey did not include flights out of San Francisco International (SFO) and Los Angeles International (LAX). This, of course, reduces the accuracy of the estimation based on the remaining routes, for San Francisco International and Los Angeles International are by far the most important airports in the corridor. However, because the calibration technique uses a sample of travel records randomly selected from the trip file, it can be said that the loss of accuracy in the analysis is only to the extent that the sample used may be considered biased.

Inventory data including information on the socioeconomic characteristics of the study area and its population and information describing the air transport system in the study area were also acquired. The socioeconomic characteristics included were population, income characteristics, and employment levels, and the transportation variables were (a) schedule frequencies of service between airport pairs, (b) line-haul travel times between airport pairs, (c) air fares between airport pairs, and (d) ground access times between population centers and airports.

The socioeconomic variables were obtained from the 1970 census (10). The transportation characteristics were obtained from available sources such as the Official Airline Guide and available road maps of the study area.

MODEL CALIBRATION AND TESTING

Air Travel Generation Model

Multiple regression analysis was performed on the logarithmic forms shown in Eqs. 1, 2, and 3. The results of this analysis for both business and nonbusiness travel are given in Table 1, from which some interesting observations can be made.

1. In all regressions, population and median income seemed to be highly significant in explaining total travel generations. The positive signs of the elasticities were as expected.
2. In the case of business travel, shortest travel time t_{ij} did not seem to be so highly significant as the other variables, even though the parameters associated with it were all negative, as expected. This is probably due to the fact that there is very little variation in this variable among the zone pairs in the study corridor. In the case of nonbusiness travel, this variable is not significant. This result seems intuitively appealing inasmuch as it is reasonable to deduce that nonbusiness travelers, i.e., mainly recreational travelers, are not sensitive to travel time.
3. In all models tested, the total explanatory power was rather low. R^2 values fall in the range 0.26 to 0.36. Because the explanatory power of the variables in the models seemed sufficiently high as explained earlier, it seems likely that additional variables describing the socioeconomic nature of the various cities in the corridor should have been included.

Based mainly on these results, it was concluded that the models as calibrated were not suitable for forecasting travel demand. On the other hand, the explanatory power of the variables included in the model seemed sufficiently high to warrant use of the models. Because the demand elasticities of variables such as population and income were estimated with sufficiently high confidence, it should be possible to use them in relating changes in income and population to changes in travel demand.

The general structure of the travel generation model is

$$T_{ij} = \prod_{k} X_k^{\alpha_k}$$

where α_k is the elasticity of the travel demand with respect to variable X_k, the ratio of relative changes of T and X, and is given by

$$\alpha_k = \frac{\partial T_{ij}/T_{ij}}{\partial X_k/X_k} \tag{17}$$

for all k. The total relative change in T_{ij} that is brought about by changes in variable X_k can be calculated from the equation for the total derivative as follows:

$$dT_{ij} = \sum_k \frac{\partial T_{ij}}{\partial X_k} dx_k$$

from which

$$\frac{dT_{ij}}{T_{ij}} = \sum_k \alpha_k \frac{dX_k}{X_k} \tag{18}$$

For example, in model B4 (Table 1), if both population of the origin city and average median income simultaneously increase by 10 percent, then the total increases in travel generation will be as given in Eq. 18.

$$\frac{dT_{ij}}{T_{ij}} = 0.29\,\frac{dP_i}{P_i} + 0.89\,\frac{dY_{ij}}{Y_{ij}} = 11.80\%$$

This procedure is used to apply growth rates to the actual city pair volumes rather than to volumes obtained from the regression model. This avoids the forecasting difficulties caused by the weak explanatory power of the regression model.

The Choice Model

To determine the probabilities P[ijk] as shown in Eq. 13 required that the distribution functions $f_1(\alpha)$, $f_2(\beta)$, and $f_3(\gamma)$ be estimated. To do this we subdivided the data into randomly selected groups. For each group, estimates of α, β, and γ were obtained by regressing on the function:

$$T_{ijk} = F^{\alpha}_{ijk}\, C^{\beta}_{ijk}\, H^{\gamma}_{ijk} \tag{19}$$

where H_{ijk} is the access time at both ends of a trip between i and j by route k. The particular form of Eq. 19 was selected from a number of alternatives that were tested statistically.

This procedure is analogous to selecting random observations on the values of α, β, and γ. Although the sample subgroups were selected at random, there is no evidence that they do represent homogeneous subsets of the population and that the readings obtained for α, β, and γ are truly disaggregate estimates. On the other hand, this procedure provides a closer approximation to a completely disaggregate model than a deterministic model.

Estimated values for the three parameters were obtained for both business and nonbusiness travel. Both α and β have the correct sign. The parameter γ does not seem to have a consistent sign; however, the F-statistic associated with this parameter is very low in all cases, indicating that it is not significantly different from zero. This is not surprising for a number of reasons. First, it was found through an investigation of the data base that access time variations between the different trip data records were not very large. Second, when compared with the effect of schedule frequency, the access time effect seemed dwarfed. Variations in schedule frequencies among airport pairs were such that the resulting variations in expected schedule delays would be considerably larger than differences among access times.

The overall statistical goodness of fit was demonstrated by the high values of coefficients of multiple determination R^2, which were over 0.90 in all cases. These were corroborated by low values of the standard error of estimate—between 0.23 to 0.39.

Table 1. Results of regressions.

Variables	Business Models B1	B2	B3	B4	B5	Nonbusiness Models NB1	NB2	NB3
Constant	-6.2	-7.97	-4.39	-7.32	-13.44	-16.10	-15.65	-15.65
$\ell n\ P_i$	0.31 (36.88)	0.29 (32.98)	0.31 (34.46)	0.29 (32.62)	0.30 (35.62)	0.32 (23.88)	0.31 (22.67	0.31 (22.67)
$\ell n\ P_j$	0.31 (36.00)	0.37 (63.22)	0.34 (62.99)	0.37 (61.23)	0.40 (99.57)	0.40 (42.42)	0.42 (56.94)	0.42 (56.94)
$\ell n\ Y_i$	1.02 (11.60)					1.05 (7.10)		
$\ell n\ Y_j$	-0.12 (0.15)					0.45 (1.49)		
$\ell n\ Y_{ij}$[a]		0.48 (4.70)	0.63 (2.23)	0.89 (4.20)	1.10 (7.50)		1.40 (7.37)	1.40 (7.37)
$\ell n\ t_{ij}$	-0.46 (3.24)	-0.32 (1.54)	-0.41 (2.27)	-0.33 (1.70)		0.18 (0.32)	0.24 (0.63)	
$\ell n\ LS_{ij}$					0.31 (0.99)			
R^2	0.3279	0.3128	0.3074	0.3117	0.3101	0.2619	0.3576	0.3558

Note: Numbers in parentheses are F-statistics.

[a]In model B2, $Y_{ij} = Y_i \times Y_j$; in model B3, $Y_{ij} = (P_i \times Y_i + P_j \times Y_j)/(P_i + P_j)$; in other models, $Y_{ij} = (Y_i + Y_j)/2$.

Figure 3. Cumulative distribution of departure frequency elasticity for business travelers.

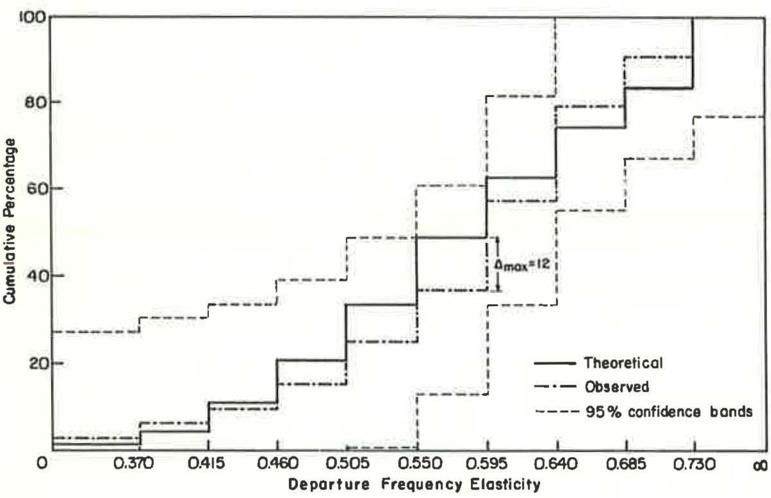


Table 2. Results of distributions and chi-square tests.

Trip Purpose	Variable	Gamma Distribution K	λ	Γ(K)	Normal Distribution μ	σ	Degrees of Freedom	χ^2 Calculated	P-Value
Business	α	23.72	38.96	1.05×10^{22}	—	—	5	6.042	0.3019
	β	4.28	14.43	8.63	—	—	4	3.925	0.4170
	γ	—	—	—	0.0136	0.1068	3	3.244	0.3592
Nonbusiness	α	4.64	13.43	14.13	—	—	2	1.596	0.4531
	β	5.33	10.69	40.19	—	—	2	1.033	0.5964
	γ	—	—	—	0.363	0.247	3	2.650	0.4520

Note: Gamma distribution: $f(x) = \frac{\lambda^K}{\Gamma(K)} x^{K-1} e^{-\lambda x}$. Normal distribution: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$

The next step was to estimate the density distribution of each of the estimates based on the values obtained in the regressions. This was done for α, β, and γ, in spite of the fact that γ was previously judged not significant. This allowed the investigation of any effect, regardless of its significance, of access time on the choice process. Furthermore, by including all parameters in the analysis, we could develop a process that is general enough to be used under other empirical conditions. This will only allow the corroboration of the rather limited results of this study.

Estimation of the density distribution functions of parameters α, β, and γ was performed by inspecting their graphical representations and then testing the fit to postulated statistical distribution functions. There is no obvious relationship between behavioral assumptions and specific statistical distribution functions. At this stage of knowledge regarding the behavioral implications of stochastic aggregation in travel demand models, the best that can be done is empirical analysis.

Graphical representations of the empirical distributions of α, β, and γ were obtained by constructing cumulative histograms for each parameter. An example of these histograms is shown in Figure 3 together with the theoretical distribution and the 95 percent confidence band. γ distributions were postulated for the parameters α and β, whereas a normal distribution was postulated for γ in both the business and nonbusiness cases. After the parameters of those hypothesized distributions were estimated from the respective data sets, statistical tests of goodness of fit were performed. Chi-square tests were performed on all six distributions and had high P-values, showing in all cases that the empirical distributions and the theoretical distributions were not significantly different.

To corroborate the results of the chi-square tests and to remove any doubt that may be precipitated because of the chi-square test's sensitivity to small sample sizes, we conducted Kolmogorov's D-test. This D-test result is shown in Figure 3 in the form of the 95 percent confidence band. As can be seen, the theoretical distribution falls within this band; therefore the postulated distribution is a valid representation of this random variable. The equations for the theoretical distributions as well as the results of the chi-square tests are given in Table 2. The assumption of the independence of α, β, and γ was checked by calculating the correlation coefficients. These were on the order of 0.3 to 0.4, which is significantly low for the sample sizes in question.

The final step in the calibration of the choice model is to evaluate the three-dimensional integral of Eq. 13. It was not possible to evaluate the integration analytically. However, it is always possible to evaluate a finite integral numerically with the aid of a high-speed computer. It is easy to tell from inspection of the integrand

$$\frac{F_{ijk}^{\alpha}\ C_{ijk}^{\beta}\ H_{ijk}^{\gamma}}{\sum_{r} F_{ijr}^{\alpha}\ C_{ijr}^{\beta}\ H_{ijr}^{\gamma}}\ f_1(\alpha)f_2(\beta)f_3(\gamma)$$

that it is indeed finite. The first part of the integrand is a ratio known to be less than unity and the second part is the joint density functions of three random variables that are also limited to unity.

The numerical analysis consisted of inputting characteristics of the 12 alternative routes in the study corridor and operating the model in an attempt to reproduce the observed data.

The overall goodness of fit of model results was then tested. Figures 4 and 5 show comparisons of model results with observed data for business and nonbusiness travel. Although a perfect fit was not achieved, in view of the results presented above and the imperfections of the data base used in calibrating the choice model, model results can generally be considered good and the calibrated model can be used for making travel forecasts.

DEMAND FORECASTING AND SENSITIVITY ANALYSIS

The first step in performing demand forecasting for STOL transportation is to postulate STOL system characteristics. Two basic assumptions are implicit in this approach: (a) that any transportation service can be represented by a number of attributes associated with it and that the decision process by which travelers choose among alternative services is essentially unchanged by the introduction of STOL or any other transportation service and (b) that the traveler's decision process remains unchanged over time. In other words, the values of the parameters and elasticities that reflect the traveler's response to exogenous influences will not change over the forecasting period. This latter assumption can be validated only after repeated applications of the forecasting models at different points in time.

STOL System Configurations

The specifications of STOL system configurations consist of the locations of STOLports, frequencies of service, travel costs, and travel times involved. There is a lack of precise data on STOLport locations and STOL aircraft characteristics. Therefore, the system variables are treated parametrically; i.e., a number of reasonable configurations are postulated and the resulting forecasts are presented. The purpose of this type of analysis is to demonstrate the use of the forecasting models and provide a procedure by which the demand potential of alternative STOL systems can be compared.

The only locations for STOLports considered in this study are existing military fields and general aviation fields. It is believed that such airports, by the mere fact of their existence, would be the first candidates for the introduction of STOL air transportation into any urban area. In the San Francisco Bay area, candidate airports include Crissy Field, Berkeley Marina, Concord Buchanan Field, and Palo Alto Airport, and, in the Los Angeles area, they include Hawthorne Airport, Fullerton Airport, Compton Airport, and Santa Monica Airport.

In the analysis, many configurations can be generated by selecting various airports from these two groups. In this presentation we show the results for only two configurations.

Postulate STOL fares were calculated from the formula

$$\text{Fare} = \frac{\text{total cost per available seat-mile} \times \text{stage length}}{\text{load factor}} + \text{tax}$$

The range of total cost per available seat-mile was taken as 2 to 4 cents for a stage length of 400 miles, which is an average range anticipated for STOL aircraft (7). The load factor range was 0.5 to 0.7.

The frequency of service was allowed to vary in two manners. First STOL service frequency was increased from 0 to 49 weekly flights, without adjusting the frequency of service of the CTOL airport pairs. Then it was postulated that some CTOL service will essentially be replaced by STOL service, so the increase in STOL frequency was accompanied by an equal decrease in CTOL frequency.

Forecasting STOL Market Share

The first model application consisted of varying STOL fares and departure frequencies without adjustment to CTOL frequency. For the STOL system chosen, Figure 6 shows its market shares of business and nonbusiness travel and the increase in STOL market share brought about by increasing service frequencies as well as decreasing the fare. Comparing the results for business and nonbusiness travel shows that business travel is more sensitive to departure frequency than nonbusiness travel; the curves for the former are steeper. Also, comparing the distances between the curves for different fares shows that nonbusiness travel is more sensitive to fare than business travel. In both cases, the market share for STOL does not exceed 8 percent of the total.

The next step in the analysis was to introduce adjustments in the CTOL schedule frequency simultaneous to increases in STOL frequencies. This was done in two manners. First, reductions in total CTOL frequencies ranging from 10 to 90 percent were obtained by distributing these flights equally among STOL routes in the configuration studied. Second, CTOL frequencies were reduced only at routes involving either SFO or LAX or both by switching flights to STOL and distributing them among STOL routes

Figure 4. Comparison of modeled and observed business trips on each route.

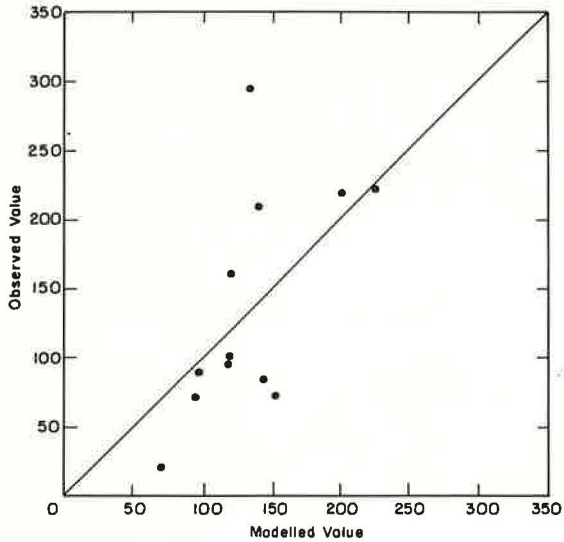


Figure 5. Comparison of modeled and observed nonbusiness trips on each route.

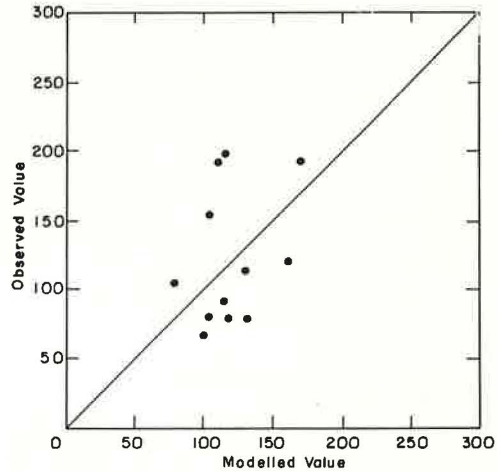


Figure 6. Sensitivity of STOL market share to departure frequency (configuration I).

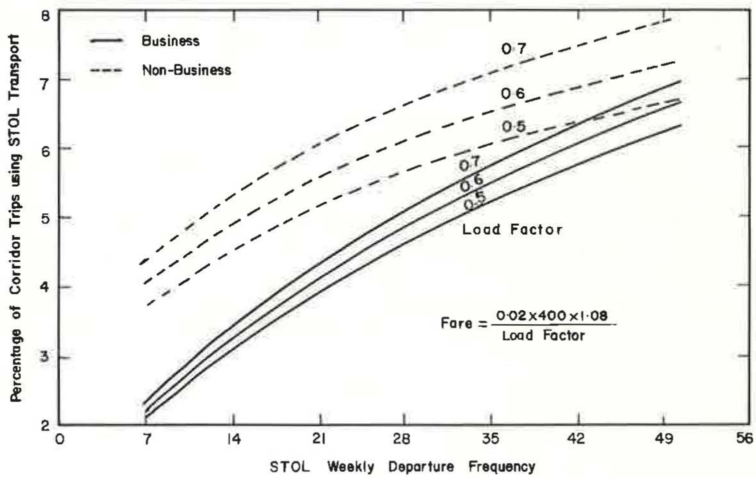
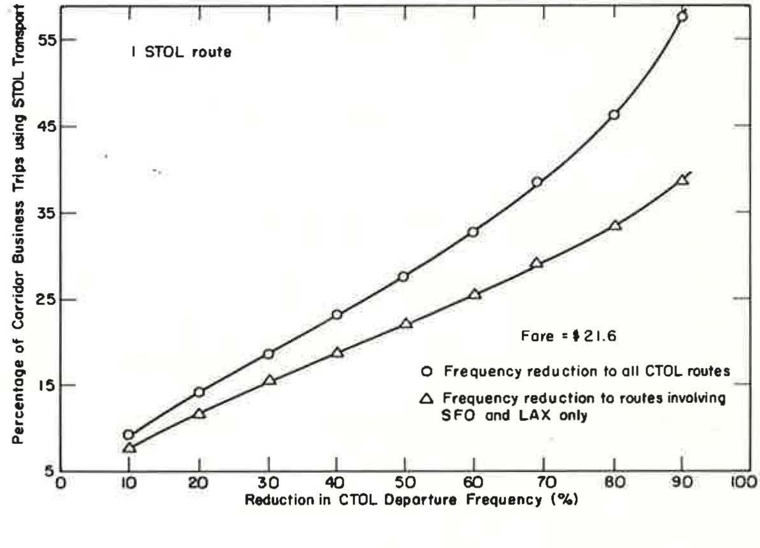


Figure 7. Sensitivity of STOL share of business travel market to changes in CTOL departure frequency (configuration I).



in the same manner as before. This second case was motivated by the idea that STOL service may be introduced to reduce congestion at major hub airports. Because only SFO and LAX may have volumes sufficiently high to cause congestion, it was assumed that reductions in CTOL service may be warranted at routes including either or both of these airports.

The results of this analysis are shown in Figure 7 for configuration I and in Figure 8 for configuration II. The figures show the increase in STOL market share related to the two types of CTOL frequency adjustments described. In configuration I a market share of more than 50 percent can be achieved; market share potential increases to a maximum of about 70 percent for configuration II. It should be noted that for both configurations the increase in business travel is larger than the increase in nonbusiness travel. This result follows from the fact that business travel is more sensitive to service frequency.

An interesting result is obtained when Figures 7 and 8 are compared. In spite of the fact that in both cases the number of flights switched from CTOL to STOL service is the same, the market share potential under configuration II is larger than under configuration I. This seems to indicate that market share increases as the number of STOL routes increases, even if the same service frequency is offered. Of course, this effect is due for the most part to the fact that a larger number of STOLports will yield a higher accessibility to STOL services.

The results obtained from applying the model to additional configurations indicated that the marginal increase in STOL market share decreases as the number of STOL routes increases. A result such as this is of vital importance when the cost-effectiveness of introducing additional STOL routes or STOLports into an urban area is analyzed.

Forecasting Total Air Travel

As was discussed, calibration results showed that the models were not sufficient to forecast the absolute levels of traffic between city pairs. However, the elasticities of demand with respect to the population, income, and travel time variables were estimated with high reliabilities. Therefore, they were used to relate the increase in travel volumes to varying growth rates in population and income and to the changes in travel times caused by the introduction of STOLports in the study area.

Based on the calibration results, the models selected were

$$\ln T_{ij} = -7.32 + 0.29 \ln P_i + 0.37 \ln P_j + 0.89 \ln Y_{ij} - 0.33 \ln t_{ij}$$

for business travel and

$$\ln T_{ij} = -15.65 + 0.31 \ln P_i + 0.42 \ln P_j + 1.40 \ln Y_{ij}$$

for nonbusiness travel. If we assume that population and income growth occurs in the same manner in all zones, simplifying Eq. 18 gives

$$\frac{\Delta T_{ij}}{T_{ij}} = (\alpha_1 + \alpha_2)W_p + \alpha_3 W_y + \alpha_4 W_t \tag{20}$$

where

α_k = elasticity with respect to variable k, and
W_k = proportional change in variable k.

If $\Delta T_{ij}/T_{ij}$ is denoted by β and the number of years over which the forecast is performed by N, future traffic volumes T^*_{ij} can be obtained from present volume T_{ij} by

$$T^*_{ij} = (1 + \beta)^N T_{ij} \tag{21}$$

The total corridor travel T^* at year N is then

Figure 8. Sensitivity of STOL share of business travel market to changes in CTOL departure frequency (configuration II).

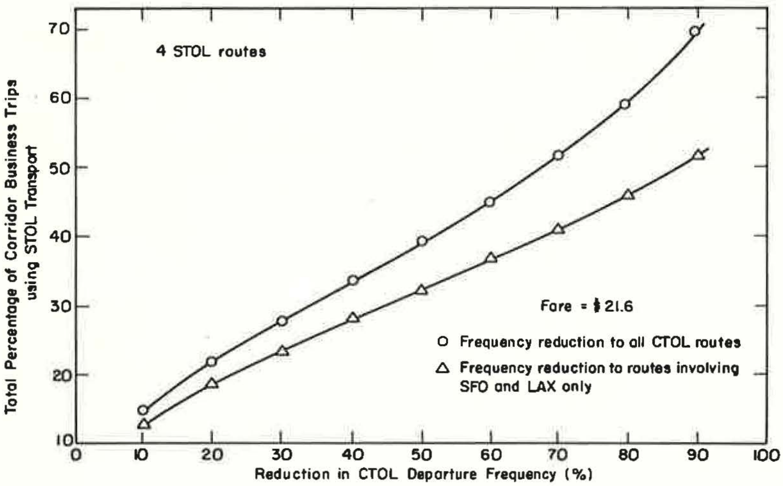


Figure 9. Sensitivity of business travel to changes in population and median family income.

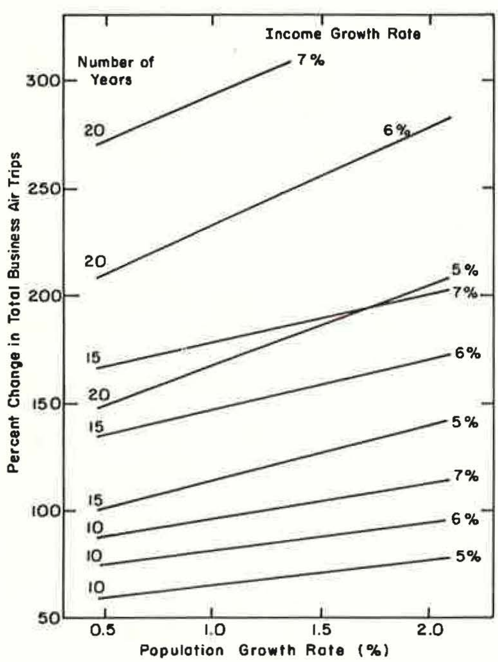
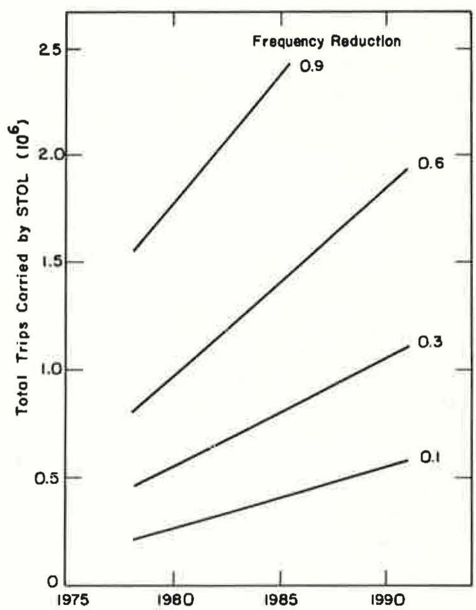


Figure 10. Forecast of STOL transport demand generated by configuration I with frequency reduction of all CTOL routes.



$$T^* = \sum_{ij} T_{ij}(1 + \beta)^N = T(1 + \beta)^N \tag{22}$$

This procedure relates future travel volumes in each city pair to present volumes and thus avoids zone-by-zone errors that may be introduced if the volume levels are forecast directly from the model.

Annual population growth rates were varied from 0.5 to 2.0 percent. Median income was increased in the range 5.0 to 7.0 percent per year. The forecast was performed for values of N of 10, 15, and 20 years. For the STOL system configuration, the following assumptions are made. During the first 10 years, i.e., up to the year 1980, no service will be introduced at any of the STOLports. In 1980 service will be introduced according to configuration I. Travel times will then be modified but held unchanged throughout the rest of the forecasting period.

Results of model application to business travel are shown in Figure 9. It should be mentioned that these results are samples of the types of results that can be obtained from the application of the travel generation model. This application allows the estimation of the increase in total corridor air travel, as well as particular city pair volumes, under different population and income growth assumptions and for different air transport system alternatives.

Forecasting STOL Demand Potential

Forecasting STOL demand is done by combining the forecasts of the total corridor air travel demand with the forecasts of the STOL market share. This is a simple operation consisting of the multiplication of the STOL share and the total air travel volume. As an example, the forecast for configuration I was obtained, for business travel, for various levels of frequency switch from the CTOL airports to the STOLports. The forecast results (Fig. 10) are based on a population growth rate of 0.5 percent per year and a median income increase of 7 percent per year with a STOL fare of $21.60. The forecast extends from a 1970 base year total volume of 3.1 million passengers to 1990. Naturally, the validity of a forecast through 1990 depends on the validity of the assumed growth rates for population and income. These growth rates could be modified at intervals within the forecast period if this is deemed necessary.

CONCLUSIONS

The procedure presented in this paper is aimed at forecasting the demand potential for transport systems in short-haul air transportation. Particularly, the objective was to study market potentials for various STOL system configurations in a short-haul corridor, such as the Los Angeles-San Francisco corridor.

The framework used in forecasting consists of three stages: Forecast the total air travel demand in the corridor; estimate the market share for any given STOL configuration; and combine these two into a forecast of the market share for STOL. An aggregative choice model is developed for this purpose. This model is stochastic in nature and permits the aggregation of individual choice decisions across a study population. Because of the lack of suitable data, it was not possible to perform this aggregation strictly on an individual traveler basis. Therefore, it was necessary to perform the aggregation on small population subgroups, chosen at random. It is believed that such a procedure, though not strictly an aggregative procedure, is a step in the right direction, particularly based on the large amount of data required for calibrating a model to account for differences among all individuals in a population.

In the forecasting model used in this study, it is assumed that no significant changes will occur in high-speed ground transportation in the study corridor and that the air travel market and the ground travel market are essentially independent. This, of course, will not be true if technological changes occur that create competition in the corridor between air and high-speed ground transportation. Therefore, it is essential to note that the validity of the forecasts obtained by the procedure developed in this study is conditional on this assumption.

The uncertainties inherent in forecasting socioeconomic indicators such as income and population are accounted for by forecasting in a sensitivity manner. That is, the forecasts are provided for ranges in growth rates for such variables. In a long-range planning situation it is always prudent to revise such forecasting inputs and modify the forecasts if necessary.

An interesting finding of this analysis is that the demand potential for STOL transportation in a corridor served by CTOL airports is strongly dependent on the level to which corridor traffic is diverted from the CTOL airports to STOLports. This is due to the strong impact of schedule frequency on the attractiveness of any particular airport pair and the initial frequency advantage that the large CTOL airports have. It was also found in the analysis that adding STOLport pairs in the system increases the market share but at a decreasing rate. This is a finding that would be important in assessing the cost-effectiveness of introducing STOL service in a short-haul air travel corridor.

ACKNOWLEDGMENT

The research presented in this paper was supported by a grant from the NASA-Ames Research Center, Moffet Field, California.

REFERENCES

1. Statewide Master Plan of Aviation, State of California: Interim Progress Report. Daniel, Mann, Johnson and Mendenhall, 1971.
2. Howrey, P. On the Choice of Forecasting Models for Air Travel. Journal of Regional Sciences, Vol. 9, No. 2, 1969.
3. Kanafani, A. An Aggregative Model of Trip Making. Transportation Research, Vol. 6, No. 2, June 1972, pp. 119-124.
4. Lansing, J. B., and Blood, D. W. Mode Choice in Intercity Travel: A Multivariate Statistical Analysis. Survey Research Center, University of Michigan, 1964.
5. Lansing, J. B., and Suits, D. B. An Analysis of Interurban Air Travel. Quarterly Journal of Economics, Feb. 1961, pp. 87-95.
6. Lave, L. B. Demand for Intercity Passenger Transportation. Journal of Regional Sciences, Vol. 12, No. 1, 1972.
7. Technical and Economic Evaluation of Aircraft for Intercity Short-Haul Transportation. McDonnell Aircraft Corporation, St. Louis, Vols. 1-3, April 1966.
8. Metzger, W. L. An Analysis of Intercity Passenger Traffic Movement Within the California Corridor Through 1980. Stanford Research Institute, April 1966.
9. Quandt, R. E., and Baumol, W. T. The Abstract Mode Model: Theory and Measurement. In Studies in Travel Demand, Vol. II, Mathematica, September 1966.
10. Census of Population for California. U.S. Bureau of the Census, 1970.
11. Warner, S. L. Stochastic Choice of Modes in Urban Travel: A Study in Binary Choice. Northwestern University Press, 1962.

DETERMINATION OF FUNCTIONAL SUBREGIONS WITHIN AN URBAN AREA FOR TRANSPORTATION PLANNING

Thomas F. Golob and Salvatore J. Hepper,
Transportation and Urban Analysis Department,
General Motors Research Laboratories; and
John J. Pershing, Jr.*, Fairfax County Planning Commission

It is axiomatic that large urban areas are not spatially homogeneous with respect to transportation demand, supply, and impact phenomena. This paper addresses this heterogeneity in terms of the transportation planning process. A technique for using areawide travel, land use, and population data to divide an urban area into a set of functional subregions is presented. Each subregion represents a planning area, and interregion planning is proposed on a different scale of analysis. The technique is based on the statistical decomposition of origin-destination flow matrices. The decomposition method can be considered a generalized type of factor analysis in which raw data observations are used as opposed to variable correlations. The units can be any spatial aggregation of people and activities, such as census tracts or minor civil divisions, and the travel can be trips for any specific purpose or a composite of all trips. Selection in both cases depends on the objectives of the planning process. Multiple discriminant and regression analyses are then used to define the subregions in terms of differences in population and land use characteristics. Results from an application of the technique in the Detroit area are presented as a case study. Six subregions, composed of groups of minor civil divisions and central city subcommunities, were found and successfully described for home-based work travel in this urban area. The results support urban economic theories of a central city core area, suburban industrial centers, and satellite cities.

•THIS PAPER presents a specific transportation planning technique as a possible addition to the set of planning tools available to decision-makers. This technique breaks down a metropolitan area or reasonably extensive planning area into functional subregions on the basis of people's trip-making behavior. The technique can be categorized as an inductive search for hypotheses of transportation and urban form through the identification of regularities in spatially aggregated data.

Analyzing urban transportation needs and evaluating alternatives on a scale smaller than an entire urban area are addressed. Studies of activity center distributions, feeder transit services (e.g., involving dial-a-bus alternatives), and local and collector roadway systems are just a few cases in which analysis could be accomplished on a subregional basis with appropriate aggregated subregional interactions. For example, in the case of the home-based work trip, a transportation system must be able to deal with the travel needs of persons who work outside of the central business district. This paper indicates how suburban areas can be organized into meaningful subsystems for transportation analysis and planning.

The technique is based on multivariate analyses of transportation flows and related data on the characteristics of the origin and destination spatial units. The analyses

*Mr. Pershing was with General Motors Research Laboratories when this work was developed.

are all well documented in theoretical textbooks, and computer programs are readily available in standard statistical packages.

STATISTICAL METHODS

The multivariate statistical method of singular decomposition is used to determine structural similarities in a data matrix comprised of travel variables. The data are organized in the form of origin-destination (O-D) matrices used in traditional transportation planning processes. Each cell of the matrix represents a measurement of the flow of people from an origin to a destination for a common travel purpose. In the terminology of traditional multivariate data matrices, each origin spatial unit is treated as a variable and each destination spatial unit is treated as an observation. For flow matrices the metric for the variables is the same for each variable; therefore it is appropriate to decompose the data without applying standardizing transformations, such as subtracting out the mean for each variable (column in the matrix) and dividing each variable by the standard deviation. This transformation characterizes the factor analysis version of singular decomposition (2). Such standardizations result in a loss of information (both the level and dispersion of the variables in the case of factor analysis), whereas singular decomposition preserves the level of information present in the raw data.

The Eckart-Young theorem (1) and Nash's extension (4) can be used to decompose an $n \times m$ data matrix, X, into three matrices such that

$$X \cong P \Lambda Q' \tag{1}$$

where

$P = n \times r$ matrix of orthonormal (independent and normalized) vectors,
$Q = m \times r$ matrix of orthonormal (independent and normalized) vectors,
$\Lambda = r \times r$ diagonal matrix of eigenvalues (latent roots of X), and
r = number of vectors extracted ($r \leq m \leq n$).

In flow matrices, $m = n$. In the terminology of traditional factor analysis, it is convenient to let $L = (Q\Lambda)$ represent the factor loading matrix and to treat the associated vectors of L as that linear combination of the original n variables (origin spatial units) that describes each new factor. Similarly, the vectors of P represent the factor scores or evaluations of the n observations (destination spatial units) on the new latent factors.

The number of independent latent factors extracted is determined by subjective evaluation of the associated eigenvalues Λ, the first differences in eigenvalues, and the cumulative fraction of trace accounted for by each factor or eigenvalue. The fraction of trace of the cross-product ($X'X$) matrix is analogous to the percentage of variance in factor analysis, and it represents the importance of a factor in describing the information in the original data. The trade-off addressed entails a sufficient explanation of the original data (in cross-product terms) within the minimum number of factors possible.

After the number of latent factors to be extracted is determined, a rotation to simple structure is achieved by application of the successive factor varimax rotation (3).

After rotation the latent factors can be expressed in terms of the original origin variables by identifying the origin spatial units most strongly related to each factor. The strength of this relationship is proportional to the absolute value of the jth term in the ith column of the rotated factor loadings matrix. Each origin spatial unit is then assigned to the factor to which it is most strongly related. The factors are then interpreted as functional subregions with relatively homogeneous travel patterns, each subregion being delineated by the origins that are assigned to that factor.

The next phase of regional decomposition is to relate the socioeconomic characteristics (referred to as "state" variables) to the subregions and associated latent origin factors. First, a correlation analysis is employed by forming an $[(r + s) \times n]$ data matrix, z, where r = factor loading vectors, s = socioeconomic characteristic vectors, and n = origin spatial units. All correlation coefficients that are significantly different from zero at the $\alpha = 0.05$ level of significance are explored. Correlations between r and s can be used to broaden the interpretation of the factors and thus to provide a better understanding of the differences between the subregions. Such understanding can be of use in initiating detailed planning analysis for specific transportation needs.

A logical extension of the correlation analysis is the application of linear regression, in which each factor is treated as a dependent variable and the relevant socioeconomic characteristics determined from the correlation analysis are treated as the independent or explanatory variables. However, to use this approach as anything other than a broad indicator of meaningful interrelationship requires that a variable or a linear combination of variables be found that maps the socioeconomic characteristics (variables) into each factor loading of the flow matrix.

A third and more satisfying test is consequently performed to determine whether the groupings can be significantly differentiated from each other on the basis of their mean value measurements on each state variable. The multivariate analysis of variance is used to determine the significance of the overall difference among several group means on a single variable by performing an F-test on the Mahalonobis D-statistic as suggested in Tatsuoka (5). A group mean that is significantly different from other group means of a particular state variable is identified at the $\alpha = 0.01$ level of significance. A summary of these results can yield penetrating descriptions of each mutually exclusive group of origins.

The results presented later are based on the difference between groups of origin spatial units with respect to a single state variable descriptor, whereas correlation analysis and linear regression are based on the similarities of the origin spatial units as evaluated on a latent factor and one or more state variables.

A CASE STUDY APPLICATION

Data

Investigation of the interrelations between transportation behavior and urban form was based on a case study of the Detroit, Michigan, urbanized area as defined by the Bureau of the Census for the base year, 1965. The data used in this analysis were collected in 1965 during the conduct of the Detroit Region Transportation and Land Use Study (TALUS). The Southeastern Michigan Council of Governments (SEMCOG) generously provided these data to the Transportation and Urban Analysis Department of the General Motors Research Laboratories in support of research into the costs and benefits of proposed new systems of public transportation and the development of improved methods of urban transportation planning and evaluation. Two types of data have been used: home-to-work trip travel flows, which are used to determine travel patterns, and state variables, which are used to explain the socioeconomic and demographic structure of a spatially defined area.

The data were supplied in the form of observations of 1,278 transportation analysis zones as defined by TALUS within the five-county Detroit region. These regions were converted to observations on 59 central city subcommunities (CCSs) and 61 minor civil divisions (MCDs) within the urbanized area. These areas represent portions of Wayne, Oakland, and Macomb Counties. Inasmuch as each MCD was made up of one or more analysis zones, the conversion entailed aggregation by code sequences of all zones within the urbanized area. The MCDs were defined consistently with the MCDs of the Bureau of the Census, which in Michigan are incorporated municipalities or townships. Each CCS follows boundaries determined by a local census committee and delineates relatively homogeneous neighborhoods of approximately the same population.

The flow variables represent home-to-work trip movements of people from an MCD or CCS to every MCD or CCS within the urbanized area. The focus on home-based work travel is for illustrative purposes only; the technique is considered relevant for a variety of transportation issues. Gross trip flows were divided by the total number of households at each origin; the resultant measure is the number of work trips per household made to each destination. This statistic can also be considered as the probability that a household at j will make a work trip to i. The state variables consist of socioeconomic characteristics of the population by place or residence, employment characteristics by place of work, and land use characteristics of each MCD or CCS. The socioeconomic characteristics that describe the spatial units are

Median family income,
Percentage of families with incomes under $3,000,
Percentage of families with incomes over $15,000,
Percentage of households with youngest over 18,
Households with residence less than 1 year,
Percentage of households owning no car,
Percentage of households owning two cars or more,
Percentage of renter-occupied dwellings,
Households per residential acre,
Persons per occupied dwelling unit,
Percentage of household heads with less than 12 years' education,
Percentage of female household heads,
Percentage of skilled or unskilled household heads,
Percentage of nonwhite population,
Percentage of population under 20 years of age, and
Percentage of population over 64 years of age.

The types of employment included in the employment characteristics are

Professional and related,
Public administration,
Service,
Finance, insurance, and real estate,
Retail trade,
Wholesale trade,
Transportation, communication, and utilities, and
Manufacturing.

Land use characteristics of interest are the percentages of commercial, public, industrial, and recreational land.

Subregion Determination

The original 120 × 120 matrix of home-based work travel per household was reduced to seven principal components that accounted for 76.40 percent of the trace in the data cross-product matrix. This number of independent factors was determined subjectively as described earlier. A graph of the cumulative percentage of trace accounted for by each level of reduction is shown in Figure 1 for up to 15 factors; doubling the number of factors selected would have accounted for only an additional 10 percent of the original variance.

By grouping each of the 120 MCD or CCS origins according to the factor on which it loaded most highly, we delineated six functional subregions of homogeneous travel patterns. The seventh factor was found to be much weaker and only represented an orientation slightly different from the others. Dimensionality was thus reduced in making the transition from factors to subregions, which is an effective check against accepting too many factors through the extraction cutoff procedure. The six subregions are shown in Figure 2. The factors are discussed in order of the fraction of trace (percentage of original variance) accounted for after rotation.

The first factor accounted for 22.1 percent of the original trace, and the subregion determined by the factor includes almost all of the central city of Detroit, as well as adjoining suburban areas along major radial transportation corridors leading into the CBD. (The subregion is labeled central in Fig. 2.) The CBD (MCD 8 in Fig. 2) is the destination with the highest score on this factor. This suggests that central city residents tend to work within the city and that there is much less out-commuting to suburban employment centers than there is in-commuting to the central city. The four major transportation corridors of in-commuting in 1965 that can be identified with this

Figure 1. Analysis of total home-based work origins.

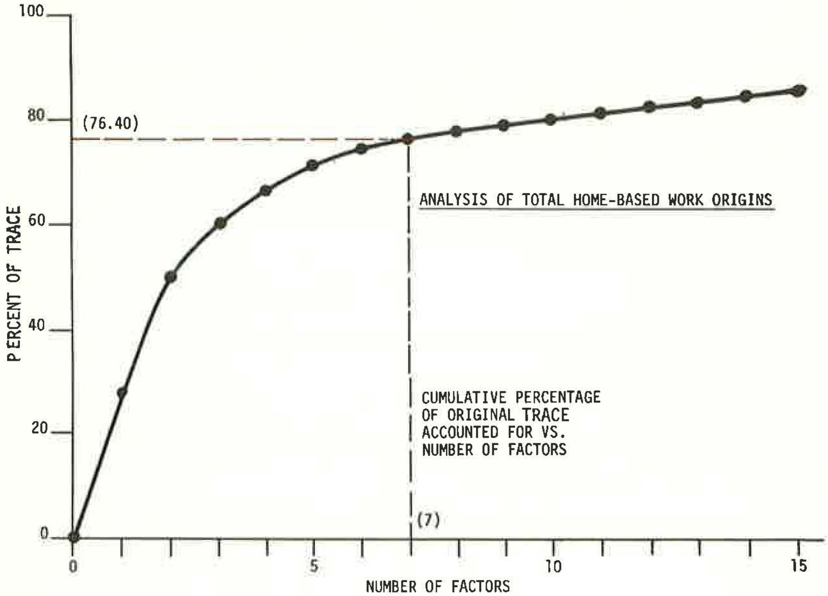
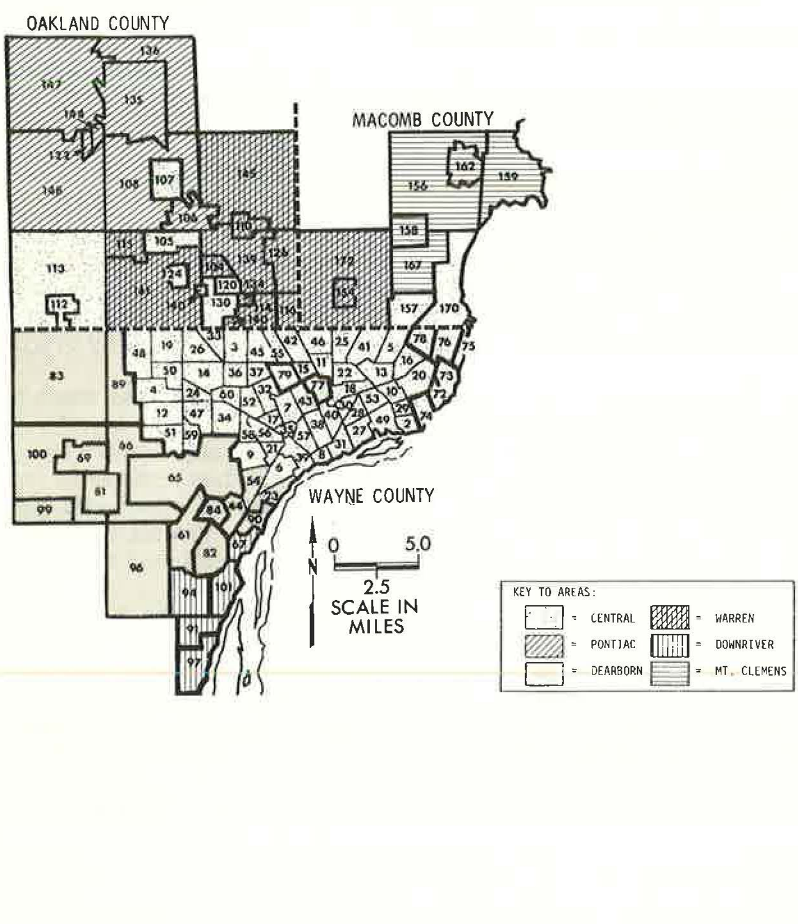


Figure 2. Groupings of MCDs by highest loadings (1965 Detroit urbanized area).



subregion were I-94/I-75 to the southwest, I-96/Grand River Avenue to the west northeast, I-75/Woodward Avenue to the north northwest, and I-94/Gratiot Avenue to the northeast.

The second factor accounted for 21.0 percent of the trace and determines the northwest portion of the urbanized area. The major destination is overwhelmingly Pontiac (MCD 135). The adjacent areas are primarily bedroom communities surrounding this industrial satellite city. The Pontiac subregion seems to be spatially separated from the central city.

Factor three accounted for 12.3 percent of the trace, and its subregion includes the southwest portion of Wayne County. Dearborn (MCD 65) is the major attractor, and Livonia (MCD 83), Ecorse (MCD 67), and Fort Wayne (MCD 23) are of secondary importance as employment attractors. Three adjacent central city subcommunities (23, 54, and 59) load highly on this factor, suggesting that the Dearborn industrial area is a more important employment center for residents of these areas than are other central city opportunities.

The fourth factor, accounting for 7.8 percent of the trace, centers on Warren (MCD 172), another major suburban industrial community. Royal Oak (139) is almost as important an attractor, whereas Ferndale (114) and Southfield (141) are somewhat lesser attractions. Again, out-commuting from the city of Detroit is not a primary characteristic in the formation of the subregion. Factor seven, which accounted for 3.6 percent of the trace, also centered on Warren but was oriented toward the east and south and did include Detroit communities; but this factor was insufficiently strong (in relation to the other six factors) to form a subregion via the origin factor loadings method used here.

Factor five accounted for only 4.9 percent of the trace, and its subregion is composed of the heavy industrial areas downriver from Detroit. Trenton (97), Ecorse (67), and Wyandotte (101) all are important suburban employment destinations represented by this factor.

The remaining sixth factor accounted for as little as 4.7 percent of the trace and centers around the satellite city of Mt. Clemens (162) to the northeast of Detroit. Warren (172), Roseville (167), and Clinton Township (156) have lesser destination scores and are thus secondary attractors for persons living in this subregion. Mt. Clemens is an established commercial center and the county seat of Macomb County.

When these six subregions are examined as separate homogeneous areas, interesting results are obtained. Between 38 and 74 percent of the home-based work trips originating within each of the six subregions have a destination also within the origin subregion. For the Detroit urbanized area as a whole, 65.3 percent of the home-based work trips remain within the origin subregion. This fact suggests that transportation planning should be oriented more toward the analysis of functional urban subregions. The within-subarea trips are given in Table 1.

Analytical Description of Subregions

The next phase of the case study dealt with the identification of the seven factors and associated six subregions in terms of the 28 state variables given earlier. It is obvious from Figure 2 that distance and travel time are major determinants of people's home/work location behavior. All of the subregional groupings are spatially contiguous with the exception of the few suburban communities with relatively large numbers of central city commuters. However, the nature of the seven factors can also be related to the characteristics of the people who live and work in each spatial area and related land use. Correlation, regression, and discrimination analyses were thus applied to the results of the decomposition.

Correlations between each factor loading and the 28 state variables with absolute values of 0.20 or more were considered in the analytical description. These results are given in Tables 2, 4, 6, 8, 10, and 11. Stepwise regression was then used to determine whether the factor loadings could be predicted from the socioeconomic characteristics in a linear manner. Functions in which all t-ratios were significant at the $\alpha = 0.05$ level were selected, up to a maximum of five variables. Problems of multicollinearity were not considered since the purpose of this portion of the study was

Table 1. Total work trips within origin subregion.

Subregion	Percentage Within Subregion
Central	74.3
Pontiac	71.5
Dearborn	56.7
Warren	47.8
Downriver	55.8
Mt. Clemens	38.0
Total urbanized area	65.3

Table 2. Significant correlations for the central subregion factor.

Variable	Correlation
Percentage of population over 64	0.47
Percentage of households with youngest over 18	0.36
Households per residential acre	0.33
Percentage of female household heads	0.33
Percentage of service employment	0.28
Percentage of renter households	0.24
Percentage of finance, insurance, and real estate employment	0.21
Percentage of family incomes under $3,000	0.20
Percentage of households with no car	0.20
Percentage of commercial land	0.20
Percentage of residence under 1 year	-0.20
Percentage of manufacturing employment	-0.22
Percentage of blue-collar household heads	-0.25
Percentage of households with two or more cars	-0.26
Persons per occupied dwelling unit	-0.35
Percentage of population under 20	-0.45

Table 3. Characteristics of central subregion.

Characteristic	Sub-region Mean	Urban Area Mean	Significantly Different Means				
			Pontiac	Dearborn	Warren	Downriver	Mt. Clemens
Persons per occupied dwelling unit	3.26	3.43		3.67	3.65	3.87	3.90
Percentage of family incomes under $3,000	17.30	14.12			7.47		
Percentage of households with youngest over 18	11.12	9.99		8.44		6.40	
Percentage of population under 20	37.51	40.15		44.28	42.64	46.94	48.24
Percentage of population over 64	10.47	8.53	5.27	5.96	6.43	4.21	4.47
Percentage of renter households	32.02	26.27			14.13		
Percentage of blue-collar household heads	33.14	36.73		46.81			53.73
Percentage of households with no car	23.42	19.53			10.06		
Percentage of households with two or more cars	24.92	28.93	40.09		36.18		
Households per residential acre	9.65	7.77	2.91		4.21		
Percentage of female household heads	19.79	16.54	11.11	12.61	10.76		
Percentage of professional employment	13.88	13.01			7.37		
Percentage of service employment	12.21	9.88		5.78			

Table 4. Significant correlations for the Pontiac subregion factor.

Variable	Correlation
Percentage of households with 2 or more cars	0.27
Households per residential acre	-0.20

Table 5. Characteristics of Pontiac subregion.

Characteristic	Sub-region Mean	Urban Area Mean	Significantly Different Means				
			Central	Dearborn	Warren	Downriver	Mt. Clemens
Percentage of family incomes over $15,000	28.73	12.60		5.23			
Percentage of blue-collar household heads	30.31	36.78					53.73

Table 6. Significant correlations for the Dearborn subregion factor.

Variable	Correlation
Percentage of blue-collar household heads	0.26
Percentage of transportation, communications, and utility employment	0.26
Percentage of population under 20	0.25
Percentage of industrial land	0.25
Persons per occupied dwelling unit	0.22
Percentage of service employment	-0.20
Percentage of family incomes over $15,000	-0.21
Percentage of population over 64	-0.26

basically exploratory. It was found that 13 of the 28 variables entered into the seven multiple regressions.

Finally, discriminant analysis was applied to the six subregional groupings of MCD/CCSs to determine whether they could be significantly discriminated from each other by the state variables. The F-ratio test was used to determine significant groups at a 1 percent level of significance. Variables that distinguish the six groupings most effectively in this manner are given in Tables 3, 5, 7, 9, and 12. The group means for each value are shown for each subregion that was significantly differentiated. Of the 28 socioeconomic variables, 16 were able to distinguish between at least one pair of the groups significantly. As expected, the central subregion proved to be the most complex. Significant correlations for the central subregion are given in Table 2. This central factor was related through multiple regression to three variables taken in combination: percentage of households with youngest child under 18 years of age (+), population over 64 years old (+), and manufacturing employment (-); a multiple R^2 of 0.55 was obtained.

The central subregion is the most easily discriminated group (Table 3). These results are generally in keeping with urban economic and sociological definitions of core areas of large metropolitan environments.

The Pontiac subregion factor loadings did not correlate very highly with the state variables. However, percentage of households with two or more cars was positively related and households per residential acre negatively related (Table 4).

Pontiac was related through multiple regression to residency for less than 1 year (+) and percentage of households with two or more cars (+), but this regression accounted for only 39 percent of the variance in this factor loading vector. Also, the Pontiac grouping was found to be distinguishable in skilled and unskilled household heads and family incomes over $15,000 (Table 5). The statistics relate to the sphere of influence of an industrial satellite city of a large metropolitan environment.

Positive correlations for the Dearborn subregion are given in Table 6. The Dearborn regression equation included population over 64 (-), percentage of industrial land (+), percentage of transportation, communications, and utilities employment (+), and percentage of professional employment (+). The regression accounted for 49 percent of the variance. The Dearborn grouping was distinguished from the other subregions by percentage of families with incomes over $15,000, percentage of professional employment, percentage of skilled and unskilled household heads, and percentage of service employment (Table 7). These results are consistent with theories of predominantly blue-collar workers with large, young families living close to an extensive suburban industrial center.

The Warren subregion factor was also rather complex, for it includes lower middle to upper class communities. Positive and negative correlations are given in Table 8.

The first Warren factor was related through regression to percentage of households with no car (-), percentage of industrial land (-), and percentage of manufacturing employment (+); the R^2 was 0.49. Important characteristics of the Warren grouping discovered through discriminant analysis were low values on families under $3,000, renter-occupied dwellings, households with no cars, and professional employment (Table 9). An important industrial suburban center similar to Dearborn is indicated, although the difference between the population characteristics of this subregion and the central area are accentuated in this case.

The Downriver subregion factor was correlated with two related indexes (Table 10). Downriver was regressed on percentage of family incomes under $3,000 (-), percentage with youngest child over 18 (-), and percentage of industrial land (+); the level of explanation was 48 percent. Discriminant analysis results pointed out the high mean value of 21.38 for the subregion grouping on percentage of industrial land. Knowledge of the area as a heavy lineal (i.e., riverfront) industrial development with adjacent blue-collar residential neighborhoods cross-validates these results.

The Mt. Clemens subregion factor was positively and negatively correlated with the variables given in Table 11. The factor was jointly related to family incomes under $3,000 (+), percentage of skilled and unskilled household heads (+), households with no car (-), and employment in public administration (+); the multiple R^2 was 0.47.

Table 7. Characteristics of Dearborn subregion.

Characteristic	Sub-region Mean	Urban Area Mean	Significantly Different Means				
			Central	Pontiac	Warren	Downriver	Mt. Clemens
Percentage of family incomes over $15,000	5.23	12.60		28.73			
Percentage of professional employment	15.71	13.01			7.37		
Percentage of service employment	5.78	9.88	12.21				

Table 8. Significant correlations for the Warren subregion factor.

Variable	Correlation
Percentage of households with two or more cars	0.36
Median family income	0.31
Persons per occupied dwelling unit	0.26
Percentage of industrial land	-0.21
Percentage of population over 64	-0.22
Percentage of residence under 1 year	-0.23
Percentage of education under 12 years	-0.30
Households per residential acre	-0.31
Percentage of nonwhite population	-0.31
Percentage of female household heads	-0.38
Percentage of family incomes under $3,000	-0.38
Percentage of renter households	-0.40
Percentage of households with no car	-0.42

Table 9. Characteristics of Warren subregion.

Characteristic	Sub-region Mean	Urban Area Mean	Significantly Different Means				
			Central	Pontiac	Dearborn	Downriver	Mt. Clemens
Percentage of family incomes under $3,000	7.47	14.12	17.30				
Percentage of renter households	14.13	26.27	32.02				
Percentage of households with no car	10.06	19.53	23.42				
Percentage of professional employment	7.37	13.01			15.71		

Table 10. Significant correlations for the Downriver subregion factor.

Variable	Correlation
Percentage of industrial land	0.35
Percentage of manufacturing employment	0.24

Table 11. Significant correlations for the Mt. Clemens subregion factor.

Variable	Correlation
Percentage of blue-collar household heads	0.27
Percentage of public administration employment	0.26
Percentage of population under 20	0.20
Percentage of female household heads	-0.20
Percentage of households with no car	-0.21

Table 12. Characteristics of Mt. Clemens subregion.

Characteristic	Sub-region Mean	Urban Area Mean	Significantly Different Means				
			Central	Pontiac	Dearborn	Warren	Downriver
Percentage of blue-collar household heads	53.73	36.78		30.31			
Percentage of public employment	14.19	4.41	3.85	4.04	3.52	6.03	2.03

Finally, the Mt. Clemens grouping was distinguished by skilled and unskilled household heads as well as employment in public administration (Table 12). These statistics indeed describe an established commercial and light-industrial satellite city.

CONCLUSIONS AND DIRECTIONS FOR RESEARCH

This paper has presented a method of reducing a large and complex urban region into smaller subregions for transportation planning purposes. It was shown that a large matrix of origin-destination flows can be reduced from 120 columns to seven without losing a great deal of the original information. Singular decomposition enables the reduction of complex matrices into smaller, more easily handled matrices and at the same time provides for the delineation of subregions exhibiting similar travel behavior analyzed—in this case home-based work trips.

These subregions were found to account for two-thirds of the home-based work trips for the total region and could be described in terms of differences in population, employment, and land use characteristics. More broadly, the subregions can be related to urban economic theories of activity distribution within metropolitan environments.

Ultimate tests of the validity of the outputs of the technique and its usefulness in urban transportation planning will have to be deferred until it can be tested in practice. One possible test might be to compare subregion results from application of the technique with results from consensus judgments of a panel of experts intimately involved with a particular planning question.

This study also makes explicit the importance of other factors besides distance or travel time in the analysis of trip-making behavior. The characteristics of the household, the employment centers, and the overall land use patterns also play an important part in travel behavior. Consequently, several areas of analysis opened up by this study should be pursued further.

The importance of distance and travel time in travel behavior needs to be considered in a subregional and regional context. What is the importance of multiemployment centers and their relationship to existing and potential transportation systems? How is time interrelated with the socioeconomic characteristics of the population, the distribution of jobs by types, and the patterns of land use?

The technique should be extended to the study of various motivations and means of trip-making. Likewise, the analysis should consider other trip purposes, such as shopping and social and recreational travel. Similarly, complex trips that involve modal splits and are multipurpose should eventually be handled. But these extensions can be deferred until the results of more simplistic applications, such as the one reported here, can be competently and objectively evaluated.

REFERENCES

1. Eckart, C., and Young, G. The Approximation of One Matrix by Another of Lower Rank. Psychometrika, Vol. 33, 1936, pp. 73-113.
2. Harman, H. H. Modern Factor Analysis. University of Chicago Press, 1967.
3. Horst, P. Factor Analysis of Data Matrices. Holt, Rinehart, and Winston, 1965.
4. Nash, D. The Generalized Eckart-Young Principal Axis Theorem. General Motors Research Laboratories, 1972.
5. Tatsuoka, M. M. Multivariate Analysis: Techniques for Educational and Psychological Research. John Wiley and Sons, 1971.

ALTERNATIVE TRAVEL BEHAVIOR STRUCTURES

STRUCTURE OF PASSENGER TRAVEL DEMAND MODELS

Moshe E. Ben-Akiva, Transportation Systems Division,
Massachusetts Institute of Technology

This study is concerned with the structure of travel demand models. Two alternative structures are defined, simultaneous and recursive, that are based on different hypotheses about the underlying travel decision-making process. The simultaneous structure is very general and does not require any specific assumptions. The recursive structure represents a specific conditional decision structure, i.e., the traveler is assumed to decompose his trip decision into several stages. Thus, simultaneous and recursive structures represent simultaneous and sequential decision-making processes. Theoretical reasoning indicates that the simultaneous structure is more sensible. Moreover, if a sequence assumption is accepted, there are several conceivable sequences, and generally there are no a priori reasons to justify a selection among them. A simultaneous model, however, is very complex because of the large number of alternatives that a traveler faces in making his trip decision. An empirical study is conducted to investigate the feasibility of a simultaneous model and to appraise the sensitivity of predictions made by a travel demand model to the structure of the model. The data set for the study was drawn from conventional urban transportation study data. Included in a trip decision are destination and mode choices. With the same data set, three disaggregate probabilistic models are estimated for the shopping trip purpose: a simultaneous model and two recursive models with two possible sequences. The simultaneous model proved to be feasible in terms of the computational costs and the estimation results. The results of the recursive models showed that estimated model coefficients vary considerably with different model structures. The simultaneous model structure is recommended.

•DECISION-MAKING in transportation planning, as in any other planning activity, requires the prediction of impacts from proposed policies. One of the inputs to the prediction process is the demand function that describes consumers' expected use of transportation services.

The approach most widely used to predict passenger travel demand (6, 12, 13) is the aggregate urban transportation model system (UTMS). [A model can be expressed mathematically in many different ways. The word structure refers to the format of writing a model that has a behavioral interpretation. A model can be used for forecasting in a format that has no behavioral interpretation. The distinction between direct and indirect travel demand model (12) is based on the format used for forecasting and does not necessarily imply a different behavioral interpretation.] It is characterized by a recursive, or sequential, structure that represents a conditional decision-making process; i.e., it is assumed that the traveler makes his trip decision in several stages. A trip decision consists of several travel choices, e.g., mode and destination. In a recursive structure the travel choices are determined one at a time, in sequence.

Two recent developments in modeling travel demand have stimulated the present study. The first was the recognition that the representation of the trip decision as a sequential process is not completely realistic. It has been argued (9) that the trip decision should be modeled simultaneously with no artificial decomposition into sequential stages. Attempts to develop simultaneous models followed the conventional approach of aggregate demand analysis, in which the quantity demanded is taken as a continuous variable (5, 8, 16, 17). The second development was the introduction of disaggregate probabilistic demand models that relied on a more realistic theory of choice among qualitative trip alternatives. However, all the disaggregate models that were developed could be used either for a single stage of the UTMS (18) or, more recently, for all the stages, but again with the assumption of a recursive structure (4).

The common denominator of these two developments is clearly a disaggregate probabilistic simultaneous travel choice model. However, because of the large number of alternative trips that a traveler faces and the large number of attributes that describe each alternative, a simultaneous model can become very complex. This raises some important issues concerning the feasibility of a simultaneous model and the sensitivity of travel predictions to the simplifying assumption of a recursive structure.

The purpose of this research is to investigate these issues and to recommend a strategy for structuring travel demand models. This study explores alternative travel demand model structures and their inherent behavioral assumptions. An empirical study is conducted to calibrate the alternative models and furnish some evidence of the feasibility and desirability of disaggregate simultaneous travel choice models.

MODELS

In general, models are simplified representations of some objects or phenomena. This study deals with econometric models, i.e., mathematical relationships describing economic phenomena of observed variables and unknown but statistically estimable parameters. We use models to better understand real-world phenomena and to make decisions based on this understanding.

Travel demand models are used to aid in the evaluation of alternative policies by predicting the consequences of alternative policies or plans. A model that determines travel consequences independently of the characteristics of various policy options obviously cannot be used to evaluate those options (unless policies are, in fact, irrelevant to consequences).

Specification of a travel demand model involves some assumptions about the relationships among the variables underlying travel behavior. Predictions made by the model are conditional on the accuracy of the behavioral assumptions and, therefore, are no more valid than the assumptions.

A model can duplicate the data perfectly, but may serve no useful purpose for prediction if it represents erroneous behavioral assumptions. For example, consider a policy that will drastically change present conditions. In this case the future may not resemble the present, and simple extrapolation from present data can result in significant errors. However, if the behavioral assumptions of the model are well captured, the model will be valid under radically different conditions. It should be noted that this discussion is very general. Behavioral assumptions are a matter of degree inasmuch as there are many levels of detail at which behavior could be described. (For example, sensitivity to policies could be regarded as a gross level of behavioral assumptions.)

The requirement that models be policy-sensitive is necessary but not sufficient for planning purposes. An additional requirement is that the models be based on valid behavioral assumptions. A model could be policy-sensitive but be useless for policy analysis if it is not based on valid assumptions.

In general, it is impossible to determine the correct specification of a model from data analysis. It should be determined from theory or a priori knowledge based on experience with, and understanding of, the phenomenon to be modeled. Frequently there is no comprehensive theory that will prescribe a specific model. Moreover, important variables are often missing because of lack of data or measurement problems. There are other potential problems that involve the different kinds of data that could be used to estimate the model (e.g., time series versus cross section, attitudinal versus engineering) and the need to use a mathematical form that is amenable to a feasible statistical estimation technique.

The result is that we may have several alternative models to evaluate. Unfortunately, "in statistical inference proper, the model is never questioned.... The methods of mathematical statistics do not provide us with a means of specifying the model" (11). In other words, given several alternative models and a data set, statistical inference will not be conclusive on which model represents the "true" process. This does not say, however, that the data do not play a role in the selection among models. At various stages of an empirical analysis, some aspects of assumptions that do not agree sufficiently with the findings may be revised. More generally, accumulated past evidence from empirical studies influences the formulation of the assumptions of new efforts.

Suppose that we are faced with a choice among some alternative models that were not discarded in the course of data analysis. If these alternative models are based on different sets of assumptions, we should decide which set makes the most sense according to a priori knowledge about behavior, along with goodness-of-fit measures and statistical significance tests.

In modeling passenger travel demand, we are concerned with the trip-making behavior of individuals or households. Hence, a prerequisite to travel demand modeling is a set of assumptions that describe the process of trip-making decisions of these individuals or households. The basis for comparing different travel demand models should be the reasonableness (or the correspondence with a priori knowledge) of the behavioral assumptions of each model.

In this study we consider two travel demand model structures: simultaneous and recursive, each representing a different travel behavior assumption. We assume a priori that a simultaneous structure is appropriate. However, we also consider recursive models, in order to evaluate the significant differences between the two.

Disaggregate Models

The behavioral assumptions of a demand model take the perspective of an individual as he weighs the alternatives and makes a choice. An aggregate model based on consumers aggregated by location or socioeconomic category could be constructed. However, aggregation during the model construction phase will only cloud the actual relationships and can cause a significant loss of information (7, 14). An aggregate model that is based on averages of observations of socioeconomic types and geographic location would not necessarily represent an individual consumer's behavior, and the same relationships may not hold in another instance or another location. For planning purposes, we are concerned with the prediction of the behavior of aggregates of people. However, in principle, aggregation to a level required for forecasting can always be performed after estimation.

In urban transportation planning (UTP) studies the data are collected on the disaggregate level and aggregated to a zonal level for use in the conventional UTMS (13). Using this disaggregate data directly in disaggregate travel demand models can bring about large savings in data collection and processing costs. Because the data are used in the original disaggregate form and are not aggregated to the zonal level, a comprehensive home interview survey is not essential as is the case of conventional aggregate models. Previous work with disaggregate travel demand models (4, 18) indicates that it is a feasible modeling approach. Thus, disaggregate travel demand models have several practical advantages over aggregate models:

1. Possible reduction in data collection costs,
2. Transferability of the models from one area to another, and
3. Possibility of using the same set of models for various levels of planning.

The problem of aggregating a disaggregate model for forecasting requires more research. However, some simplified methods, such as the use of homogeneous market segments (1, 12), are available and can be used.

Choice Theory

In general, models that describe consumer behavior are based on the principle of utility maximization subject to resource constraints. Conventional consumer theory, however, is not suitable for deriving models that describe a probabilistic choice from a qualitative or discrete set of alternatives. Therefore, the travel demand models developed in this study rely on probabilistic choice theories (2, 3, 4, 10).

It is assumed that the consumer selects the alternative that maximizes his utility. The probabilistic behavior mechanism is a result of the assumption that the utilities of the alternatives are not certain but are random variables determined by a specific distribution.

The choice probability of alternative i is

$$P(i:A_t) = \text{Prob}[U_{it} \geq U_{jt}, \forall j \in A_t]$$

where

A_t = set of alternative choices available to consumer t and
U_{it} = utility of alternative i to consumer t.

The utilities are essentially indirect utility functions that are defined in theory as the maximum level of utility for given prices and income. In other words, U_{it} is a function of the variables that characterize alternative i, denoted as X_i, and of the socioeconomic variables describing consumer t, denoted as S_t. Thus, we can write

$$U_{it} = U_i(X_i, S_t)$$

The set of alternatives A_t is mutually exclusive and exhaustive such that only one alternative is chosen. The deterministic equivalent of this theory is simply a comparison of all alternatives available and selection of the alternative with the highest utility.

The mathematical form of the choice model is determined from the assumption about the distribution of the utility values. The coefficients of the utility functions are estimated with a cross section of consumers and by observations of actual choices. Therefore, the observed dependent variable has a value of zero or one. The forecast of the model is a set of probabilities for the set of alternatives.

The Multinomial Logit Model

There are a number of probabilistic choice models that are available; two of the most popular and most useful are the probit and logit models. The multinomial logit model, as described below, appears to be superior to probit because of the computational time requirements.

The logit model (2, 4) is written as follows:

$$P(i:A_t) = \frac{e^{U_i(X_i,S_t)}}{\sum_{j \in A_t} e^{U_j(X_j,S_t)}}$$

With disaggregate cross-sectional data, the logit model is estimated by using the maximum likelihood method (15).

The Travel Choices

A trip decision for a given trip purpose consists of several choices: trip frequency, destination, time of day, mode, and route. In a probabilistic choice approach we are interested in predicting the joint probability $P(f, d, h, m, r:FDHMR_t)$, which is defined as the probability that individual or household t will make a trip with frequency f to destination d during time of day h via mode m along route r. The set of alternatives $FDHMR_t$ consists of all possible combinations of frequencies, destinations, times of day, modes, and routes available to individual t.

For the purpose of presentation we consider only two travel choices: destination and mode. The set of all alternative combinations of destinations and modes is denoted as DM. (For simplicity we drop the subscript t.) We can partition this set according to destination to get the sets of alternative modes to a given destination M_d. If modes and destinations have no common attributes and the two choices are independent, then M_d is independent of d and can be written as M. However, this is an unrealistic assumption because there are many attributes, such as travel time, that are in fact characterized by all the travel choices. Therefore, it is assumed that $M_d \neq M_{d'}$. We are interested here in predicting the joint probability P(d, m:DM).

The Alternative Structures

If we assume that the two choices are independent, we write the following independent structure:

$$P(d:D) = \mathrm{Prob}[U_d \geq U_{d'},\ \forall d' \epsilon D]$$

$$P(m:M) = \mathrm{Prob}[U_m \geq U_{m'},\ \forall m' \epsilon M]$$

and

$$P(d, m:DM) = P(d:D) \times P(m:M)$$

where

D = set of alternative destinations,
M = set of alternative modes,
U_d = utility from destination d, and
U_m = utility from mode m.

(This is an unrealistic structure for travel demand, but it is presented for the purpose of comparison with other structures.)

Consider a conditional decision-making process in which, for example, destination is chosen first and then, conditional on the choice of destination, a mode is chosen. For this assumption we write the following recursive structure:

$$P(d:D) = \mathrm{Prob}[U_d \geq U_{d'},\ \forall d' \epsilon D]$$

$$P(m:M_d) = \mathrm{Prob}[U_{m|d} \geq U_{m'|d},\ \forall m' \epsilon M_d]$$

and

$$P(d, m:DM) = P(d:D) \times P(m:M_d)$$

where

M_d = set of alternative modes to destination d and
$U_{m|d}$ = utility from mode m given that destination d is chosen.

Assuming that the choice of mode is dependent on the choice of destination and vice versa, we can write the following simultaneous structure:

$$P(d:D_m) = \mathrm{Prob}[U_{d|m} \geq U_{d'|m},\ \forall d' \epsilon D_m]$$

$$P(m:M_d) = \mathrm{Prob}[U_{m|d} \geq U_{m'|d},\ \forall m' \epsilon M_d]$$

where D_m = the set of alternative destinations by mode m.

In the independent and recursive structures we predict the joint probability by multiplying the structural probabilities. However, in a simultaneous structure, the two conditional probabilities are insufficient information to predict the joint probability. Therefore, we need to estimate either a marginal probability, say P(d:D), or, directly, the joint probability. The problem with the first approach is that we need to define a

U_d where we originally specified $U_{d|m}$. The second approach requires a specification of the joint utility U_{dm}, in which the combination dm is considered as a single alternative. This approach is more logical because it corresponds with the notion of a simultaneous choice. Hence, in the simultaneous structure, we need to estimate the following choice probability:

$$P(d, m:DM) = \text{Prob}[U_{dm} \geq U_{d'm'}, \forall\, d'm' \epsilon DM]$$

Given the joint probability we can derive any desired marginal or conditional probability. For example,

$$P(m:M) = \sum_{d \epsilon D_m} P(d, m:DM)$$

and

$$P(d:D_m) = \frac{P(d, m:DM)}{P(m:M)}$$

Alternative Models

For simplicity, we write the probabilities in this section without the notation for the set of alternatives. In other words, we will write $P(d, m:DM_t)$ as $P_t(d, m)$, and $P(m:M_{dt})$ as $P_t(m|d)$.

In the prediction of joint probability $P_t(f, d, m, h, r)$, the set of alternatives consists of all possible trips or all possible combinations of frequencies, destinations, modes, times of day, and routes available to individual t. In a simultaneous structure of the logit model, this will be the definition of the set of alternatives, and the choice probability will be for an alternative f, d, m, h, r combination.

The joint probability can be written as a product of marginal and conditional probabilities:

$$P_t(f) \times P_t(d|f) \times P_t(m|f, d) \times P_t(h|f, d, m) \times P_t(r|f, d, m, h)$$

and can be written in many ways:

$$P_t(f) \times P_t(h|f) \times P_t(m|f, h) \times P_t(d|f, h, m) \times P_t(r|f, h, m, d)$$

In a recursive structure we will use a logit model for each probability separately and arrange the set of alternatives for each choice according to the sequence implied by the way we write the product. For example, the probability $P_t(m|f, d)$ is the probability of choosing mode m, when the set of alternatives consists of the modes available to individual t, to destination d at trip frequency f.

Calibrating a sequential model requires assumptions beyond the definitions of the relevant sets of alternatives for each choice. Consider, for example, the choice model for the probability $P_t(m|f, h)$. The problem is how to include in the model all the variables that for a given mode vary across destinations. Clearly, we cannot use all these variables as separate variables with their own coefficients. Therefore, we need to construct composite variables. There are many possible composition schemes. In addition there is the possibility of constructing the composite variables from several original variables together such that the trade-off among them is kept constant in all choices. For example, for an alternative destination we can define a generalized price by each mode that is a function of travel time and travel cost; then we aggregate across destinations to create a composite generalized price that is specific only to mode.

THE EMPIRICAL STUDY

The data for this study were taken from a data set prepared for the Metropolitan

Washington Council of Governments (WCOG). The data set was combined from a home interview survey conducted in 1968 by WCOG and a network (i.e., level of service) data set assembled by WCOG and R. H. Pratt Associates.

The scale and the objectives of this empirical study dictated that we use only a small subsample of the original data set for a single trip purpose, shopping. The data were kept in the disaggregate form where the observation unit is a household. This follows the assumption that the behavioral unit for a shopping trip is also a household.

Hence, the disaggregate data were exclusively drawn from conventional urban transportation study data. Specifically, trip and socioeconomic data from a home interview survey, level-of-service data from coded networks, and other user cost data customarily collected by transportation planning agencies were used.

Because our purpose is to evaluate the sensitivity of the predictions to the structure of the model, we consider in the empirical work only the joint probability of destination and mode (given that a trip is taken)—$P_t(m, d)$. We model this joint probability with three alternative structures: a simultaneous logit model that estimates this probability directly and the following two possible recursive model sequences:

$$P_t(d) \times P_t(m \mid d)$$

and

$$P_t(m) \times P_t(d \mid m)$$

where a logit model is applied to each probability separately. We also investigate alternative ways of constructing composite variables for the marginal probability.

The justification for separating destination and mode choices from other choices is as follows: The choice of time of day is assumed to be insignificant because the sample included only off-peak shopping trips. Route choice is not reported in the available data. The actual frequency is also not reported. Trips are reported for a 24-hour period. Therefore, the observed daily frequency is either 0 or 1 (and in a few cases 2). If we use an aggregate of households, this is sufficient information to compute an average frequency. For a disaggregate model the actual frequency is not available. We are forced to assume that the choices of mode and destination are independent of the actual frequency and, therefore, can be modeled separately. Note that with 0, 1 daily frequencies, $P_t(f = 1 \mid d, m) = 1$ and $P_t(f = 0 \mid d, m) = 0$.

The sample used for estimation consists of 123 household home-shop-home round trips that were selected randomly from a home interview sample in the northern corridor of Metropolitan Washington. Each household has a choice between two modes, the family car and bus, and several shopping destinations, ranging from one to eight according to the location of the household residence. It is important to note that we need to consider only alternatives that have positive choice probabilities. Therefore, a shopping location that is too far or a mode that is unsafe and consequently not feasible, or assumed to have negligible choice probability, need not be included in the set of alternatives.

The data consist of level-of-service variables by mode and destination, shopping opportunities by destination, and socioeconomic characteristics of the household. Each observation included the value of the variables for all the relevant alternatives for this household and the observed choice.

Specification of the Variables

The following list gives the definitions of the variables:

TO_{dm} = out-of-vehicle travel time to destination d by mode m (in minutes)

TI_{dm} = in-vehicle travel time to destination d by mode m (in minutes)

$\frac{C_{dm}}{INC}$ = out-of-pocket cost to destination d by mode m (in cents), divided by household income

E_d = wholesale-retail employment (number of jobs)

$DCBD_d$ = CBD specific dummy variable for destination d

$= \begin{cases} 1 \text{ for } d = \text{CBD} \\ 0 \text{ otherwise} \end{cases}$

DA_m = automobile-specific dummy variable for mode m

$= \begin{cases} 1 \text{ for } m = \text{automobile} \\ 0 \text{ for } m = \text{bus} \end{cases}$

$DINC_m$ = automobile-specific income variable for mode m

$= \begin{cases} \text{INC for } m = \text{automobile} \\ 0 \text{ for } m = \text{bus} \end{cases}$

The level-of-service variables are generic rather than mode-specific. (This would increase the number of level-of-service variables from three to six.) In this case, the marginal rates of substitution among level-of-service variables will differ for alternative modes. From a theoretical point of view it makes more sense to have equal marginal rates of substitution. The differences among modes that are not explained by the level-of-service variables included, such as differences in comfort and safety, are accounted for by the mode-specific dummy variables. This assumption has been tested (4) from an empirical point of view. A mode choice model was calibrated with mode-specific level-of-service variables, and it was found that the modal coefficients were not significantly different. This supports the a priori assumption of equal marginal rates of substitution.

The alternative models estimated are presented in terms of the log of the odds of choosing one alternative over another. That is, the models are expressed as

$$\text{Log}\ \frac{P(i)}{P(j)} = \sum_{k=1}^{K} (X_{ik} - X_{jk})\ \hat{\theta}_k$$

where

$P(i)$ = choice probability of alternative i,
X_{ik} = k th explanatory variable for alternative i, and
$\hat{\theta}_k$ = coefficient estimate of the k th explanatory variable.

The Simultaneous Model

In the simultaneous model presented below, the joint probability of destination and mode (given that a trip is made) was directly estimated. The sets of alternatives consist of combinations of mode and destination. There are from two to 16 alternatives for each observation. The results that were obtained are as follows:

$$\begin{aligned}
\text{Log}\ \frac{P(d,m)}{P(d',m')} = &\underset{(0.970)}{-1.36}\ (DA_m - DA_{m'}) \underset{(0.0202)}{-\ 0.0633}\ (TO_{dm} - TO_{d'm'}) \\
&\underset{(0.0116)}{-\ 0.0164}\ (TI_{dm} - TI_{d'm'}) \underset{(0.0216)}{-\ 0.0757}\ (C_{dm}/INC - C_{d'm'}/INC) \\
&\underset{(0.158)}{+\ 0.114}\ (DINC_m - DINC_{m'}) \underset{(0.0000875)}{+\ 0.000171}\ (EMP_d - EMP_{d'}) \\
&\underset{(0.554)}{+\ 0.316}\ (DCBD_d - DCBD_{d'})
\end{aligned} \qquad (1)$$

$L^*(0) = -277.678$
$L^*(\hat{\theta}) = -207.380$
$\rho^2 = 0.25$
$\bar{\rho}^2 = 0.25$

where

$P(d, m)$ = joint probability of choosing destination d and mode m,
$L^*(0)$ = log likelihood for 0 coefficients,
$L^*(\hat{\theta})$ = log likelihood for the estimated coefficients,
ρ^2 = coefficient of determination
$= 1 - \frac{L^*(\hat{\theta})}{L^*(0)}$, and
$\bar{\rho}^2 = \rho^2$ adjusted for degrees of freedom,

and the numbers in parentheses below the model coefficients are standard errors.

All the signs and the relative values of the coefficient estimates are as expected. The pure automobile effect, θ_{DA}, gave a minus sign; however, it should be interpreted as a transit bias only together with the coefficient of the automobile-specific income variable, which is positive. Out-of-vehicle travel time is on the order of four times more onerous than in-vehicle travel time. The standard errors of the coefficients of the automobile specific income and the CBD dummy variables are relatively large; however, they have the expected signs.

Alternative Recursive Models

Three alternative composition rules were used: weighted prices, weighted generalized price, and log of the denominator. The composite variables are defined when the estimation results are presented. In addition, there are two possible sequences:

1. $d \rightarrow m$:d followed by m, and
2. $m \rightarrow d$:m followed by d.

Hence, we estimated a total of six recursive models, three for each sequence. The estimation starts with the conditional probability, i.e., $P(m|d)$ in the first sequence and $P(d|m)$ in the second sequence. Then, the marginal probability is estimated by using the composite variables that are calculated with results from the conditional probabilities. Note that for each sequence there are one conditional probability and three marginal probabilities for the alternative composition rules.

Sequence $d \rightarrow m$: The Conditional Probability

The conditional probability presented below is the equivalent of a trip interchange modal-split model (20). The model predicts the probability of mode choice for a given destination (and given that a trip is made). The sets of alternatives consist of the bus and automobile modes for the chosen destination. The estimation results are as follows:

$$\begin{aligned} \text{Log}\,\frac{P(m|d)}{P(m'|d)} = &\underset{(1.33)}{-0.639}\,(DA_m - DA_{m'}) \underset{(0.0237)}{- 0.0515}\,(TO_{dm} - TO)_{dm'} \\ &\underset{(0.0261)}{- 0.0108}\,(TI_{dm} - TI)_{dm'} \underset{(0.0530)}{- 0.137}\,(C_{dm}/INC - C_{dm'}/INC) \\ &\underset{(0.199)}{+ 0.0490}\,(DINC_m - DINC_{m'}) \end{aligned} \qquad (2)$$

$L^*(0) = -85.257$
$L^*(DA) = -56.216$
$L^*(\hat{\theta}) = -23.033$
$\rho^2 = 0.73$
$\bar{\rho}^2 = 0.72$
$\rho^2_{DA} = 0.59$
$\bar{\rho}^2_{DA} = 0.58$

where

$P(m|d)$ = conditional probability of choosing mode m given that destination d is chosen,

$L^*(DA)$ = log likelihood for 0 coefficients except for pure automobile effect DA, and
ρ^2_{DA} = coefficient of determination in addition to the pure automobile effect.

It can be seen that all the coefficients have their expected signs. Out-of-vehicle travel time is almost five times more onerous than in-vehicle travel time. The standard errors of the coefficients of in-vehicle travel time and income are relatively large; however, the coefficients have their expected signs.

Sequence d → m: The Marginal Probability

The marginal probability of destination choice is the equivalent of a pre-modal-split distribution model. This model predicts the probability of destination choice with the mode choice indeterminate. The sets of alternatives consist of the alternative shopping destinations. Three models with the alternative composition rules were estimated for this probability, and the results are presented below for weighted prices.

$$\begin{aligned}\text{Log}\,\frac{P(d)}{P(d')} = {} & \underset{(0.0523)}{-0.0227}\,(TO^M_d - TO^M_{d'}) \underset{(0.0173)}{- 0.0374}\,(TI^M_d - TI^M_{d'}) \\ & \underset{(0.0327)}{- 0.0269}\,(C^M_d/INC - C^M_{d'}/INC) \underset{(0.0000910)}{+ 0.000130}\,(EMP_d - EMP_{d'}) \\ & \underset{(0.595)}{+ 0.638}\,(DCBD_d - DCBD_{d'})\end{aligned} \tag{3}$$

$L^*(0) = -192.421$
$L^*(\hat{\theta}) = -182.485$
$L^*_{dm}(\hat{\theta}) = -205.518$
$\rho^2 = 0.05$
$\bar{\rho}^2 = 0.04$
$\rho^2_{dm} = 0.26$

where

$P(d)$ = marginal probability of choosing destination d,

$TO^M_d = \sum_m TO_{dm} \times P(m|d)$,

$TI^M_d = \sum_m TI_{dm} \times P(m|d)$,

$C^M_d = \sum_m C_{dm} \times P(m|d)$,

$L^*_{dm}(\hat{\theta})$ = log likelihood for the joint probability, and
ρ^2_{dm} = coefficient of determination for the joint probability.

Note that $\bar{\rho}^2_{dm}$ is not computed. The reason is that the two separate models have different numbers of degrees of freedom. The results for weighted generalized prices are as follows:

$$\begin{aligned}\text{Log}\,\frac{P(d)}{P(d')} = {} & \underset{(0.0000867)}{0.000149}\,(EMP_d - EMP_{d'}) \underset{(0.510)}{+ 0.353}\,(DCBD_d - DCBD_{d'}) \\ & \underset{(0.141)}{+ 0.507}\,(GP^M_d - GP^M_{d'})\end{aligned} \tag{4}$$

$L^*(0) = -192.421$
$L^*(\hat{\theta}) = -184.866$
$L^*_{dm}(\hat{\theta}) = -207.899$

$$\rho^2 = 0.04$$
$$\bar{\rho}^2 = 0.04$$
$$\rho^2_{dm} = 0.25$$

where

$$GP_d^M = \sum_m (-0.0515 \times TO_{dm} - 0.0108 \times TI_{dm} - 0.137 \times C_{dm}/INC)\ P(m|d)$$

The results for the log of the denominator are as follows:

$$\text{Log}\ \frac{P(d)}{P(d')} = \underset{(0.0000862)}{0.000149}\ (EMP_d - EMP_{d'}) + \underset{(0.510)}{0.295}\ (DCBD_d - DCBD_{d'}) + \underset{(0.147)}{0.549}\ (\log P_d^{M*} - \log P_{d'}^M) \qquad (5)$$

$$L^*(0) = -192.421$$
$$L^*(\hat{\theta}) = -184.068$$
$$L^*_{dm}(\hat{\theta}) = -207.101$$
$$\rho^2 = 0.04$$
$$\bar{\rho}^2 = 0.04$$
$$\rho^2_{dm} = 0.25$$

where

$$P_d^M = \sum_m \exp(-0.639DA_m - 0.0515TO_{dm} - 0.0108TI_{dm} - 0.137C_{dm}/INC + 0.0490DINC_m)$$

All the models have relatively low coefficients of determination, which is attributed to the lack of more descriptive attraction data. All three models gave coefficient estimates with the expected signs. However, in Eq. 3, the coefficient of out-of-vehicle travel time is smaller than the coefficient of in-vehicle travel time, in contrast to what we would expect. The standard errors in Eq. 3 are relatively large; however, it fits the data as well as the two other models.

The model with weighted prices represents the assumption that the marginal rates of substitution among level-of-service attributes are different for different choices. The two other models assume equal rates for different choices. From a theoretical point of view, the latter assumption seems more reasonable. It is more likely that a traveler will have an identical trade-off between travel time and money cost for different travel choices rather than several of them, each being used for a different choice. The poor results from the weighted prices model support this assumption. It appears that all previous travel demand models reported in the literature have made the assumption of equal marginal rates of substitution for different choices.

Comparison of Eqs. 4 and 5 shows that there are no significant differences (2). The coefficient estimates of the CBD dummy variable have relatively large standard errors in the two models. However, the coefficients have the expected signs. The model shown in Eq. 4 is equivalent to the model developed by Charles River Associates (4).

Sequence m → d: The Conditional Probability

The conditional probability in this sequence is the equivalent of a post-modal-split trip distribution model. The model predicts the probability of destination choice for a given mode. The sects of alternatives consist of the alternative shopping destinations for the chosen mode. The estimation results of this model are as follows:

$$\text{Log}\,\frac{P(d|m)}{P(d'|m)} = \underset{(0.0380)}{-0.0610}\,(TO_{dm} - TO_{d'm}) - \underset{(0.0136)}{0.0287}\,(TI_{dm} - TI_{d'm})$$
$$- \underset{(0.0263)}{0.0470}\,(C_{dm}/INC - C_{d'm}/INC) + \underset{(0.0000899)}{0.000148}\,(EMP_d - EMP_{d'})$$
$$+ \underset{(0.548)}{0.330}\,(DCBD_d - DCBD_{d'}) \tag{6}$$

$$L^*(0) = -192.421$$
$$L^*(\hat{\theta}) = -179.680$$
$$\rho^2 = 0.07$$
$$\bar{\rho}^2 = 0.06$$

where $P(d|m)$ = conditional probability of choosing destination d given that mode m is chosen.

The signs of the coefficient estimates are as expected. Out-of-vehicle travel time is more than two times more onerous than in-vehicle travel time. The coefficient of the CBD dummy variable has the expected sign but a relatively large standard error. The goodness of fit of this model is relatively low because of the large number of alternatives and the lack of better attraction description.

Sequence m → d: The Marginal Probability

The marginal probability of mode choice is the equivalent of a trip-end modal-split model (20). This model predicts the probability of mode choice with indeterminate destination choice. The sets of alternatives include the bus and automobile modes. Again, we model this probability with the three alternative composition rules. The results that were obtained for weighted prices are as follows:

$$\text{Log}\,\frac{P(m)}{P(m')} = \underset{(1.27)}{-0.952}\,(DA_m - DA_{m'}) - \underset{(0.0204)}{0.0509}\,(TO^D_m - TO^D_{m'})$$
$$+ \underset{(0.0429)}{0.109}\,(TI^D_m - TI^D_{m'}) - \underset{(0.0725)}{0.183}\,(C^D_m/INC - C^D_{m'}/INC)$$
$$+ \underset{(0.225)}{0.293}\,(DINC_m - DINC_{m'}) \tag{7}$$

$$L^*(0) = -85.257$$
$$L^*(DA) = -56.216$$
$$L^*(\hat{\theta}) = -24.596$$
$$L^*_{dm}(\hat{\theta}) = -204.276$$
$$\rho^2 = 0.71$$
$$\bar{\rho}^2 = 0.70$$
$$\rho^2_{dm} = 0.26$$
$$\rho^2_{DA} = 0.56$$
$$\bar{\rho}^2_{DA} = 0.55$$

where

$P(m)$ = marginal probability of choosing mode m,

$TO^D_m = \sum_d TO_{dm} \times P(d|m)$,

$TI^D_m = \sum_d TI_{dm} \times P(d|m)$, and

$C^D_m = \sum_d C_{dm} \times P(d|m)$.

For weighted generalized price, the results are as follows:

$$\text{Log}\,\frac{P(m)}{P(m')} = \underset{(0.959)}{-2.07}\,(DA_m - DA_{m'}) + \underset{(0.157)}{0.117}\,(DINC_m - DINC_{m'})$$

$$+ \underset{(0.371)}{1.62}\,(GP_m^D - GP_{m'}^D) \qquad (8)$$

$$\begin{aligned} L^*(0) &= -85.257 \\ L^*(DA) &= -56.216 \\ L^*(\hat{\theta}) &= -31.039 \\ L^*_{dm}(\hat{\theta}) &= -210.719 \\ \rho^2 &= 0.64 \\ \bar{\rho}^2 &= 0.63 \\ \rho^2_{dm} &= 0.24 \\ \rho^2_{DA} &= 0.45 \\ \bar{\rho}^2_{DA} &= 0.44 \end{aligned}$$

where

$$GP_m^D = \sum_d (-0.0610TO_{dm} - 0.0287TI_{dm} - 0.0470C_{dm}/INC)P(d|m)$$

For the log of the denominator, the results are as follows:

$$\text{Log}\,\frac{P(m)}{P(m')} = \underset{(0.955)}{-1.74}\,(DA_m - DA_{m'}) + \underset{(0.168)}{0.0489}\,(DINC_m - DINC_{m'})$$

$$+ \underset{(0.303)}{1.42}\,(\log P_m^D - \log P_{m'}^D) \qquad (9)$$

$$\begin{aligned} L^*(0) &= -85.257 \\ L^*(DA) &= -56.216 \\ L^*(\hat{\theta}) &= -27.832 \\ L^*_{dm}(\hat{\theta}) &= -207.512 \\ \rho^2 &= 0.67 \\ \bar{\rho}^2 &= 0.67 \\ \rho^2_{dm} &= 0.25 \\ \rho^2_{DA} &= 0.50 \\ \bar{\rho}^2_{DA} &= 0.50 \end{aligned}$$

where

$$P_m^D = \sum_d \exp(-0.0610TO_{dm} - 0.0287TI_{dm} - 0.0470C_{dm}/INC$$
$$+ 0.000148EMP_d + 0.330DCBD_d)$$

Again, Eq. 7, the weighted prices model, gave unreasonable coefficient estimates, similar to those in Eq. 3. The two other models, Eqs. 8 and 9, gave better results. The coefficients of the income variable have the expected signs but relatively large standard errors. The model of Eq. 8 uses the same composition scheme as the model developed by CRA (4); however, this model assumes a different sequence.

Comparison of Alternative Models

The alternative models that gave reasonable coefficient estimates are given below.

Model	Method	Equation
Simultaneous	Direct estimation	1
Recursive d → m	Weighted generalized price	2, 4
	Log of the denominator	2, 5
Recursive m → d	Weighted generalized price	6, 8
	Log of the denominator	6, 9

It was not the purpose of this study to accept or reject the a priori assumption of a simultaneous decision-making process. As expected, the empirical evidence does not show which of the alternative structures, one simultaneous and two recursive, is more likely to be correct. All the models gave reasonable coefficient estimates. Furthermore, all the models gave essentially equal goodness of fit: $\rho^2 = 0.25$. The simultaneous model includes seven coefficients, whereas the recursive models included eight. This implies that the simultaneous model has a slight edge in this category, but it is certainly not a conclusive difference.

The simultaneous model that included observations with up to 16 alternatives and seven variables gave reasonable coefficient estimates. The computer cost was only slightly higher (≃20 percent) than the cost of a binary mode choice model with five variables. This indicates that a simultaneous model is feasible for the two choices of destination and mode. It also indicates that expanding the set of choices and therefore increasing the number of alternatives and variables may not be an unrealistic objective.

Comparison of the coefficient estimates of the simultaneous model with those of the estimated recursive models suggests that the predictions are sensitive to the structure of the model. This sensitivity can be demonstrated by showing some examples of the important trade-offs and elasticities. Table 1 gives the values of time implied by the different models.

Although the standard errors are relatively large, this is not atypical for estimates of value of time (19). (The estimated model coefficients that were used to compute the values of time were significantly different from zero.)

Estimated values of time from the simultaneous model are greater than those estimated from a mode choice model (given destination) and smaller than those estimated from a destination choice model (given mode).

Table 2 gives some direct elasticities of the mode choice probability. The figures in Table 2 are based on the following case:

1. Annual household income is between \$10,000 and \$12,000,
2. The probabilities of choosing bus and automobile are 0.2 and 0.8 respectively,
3. Out-of-vehicle travel times are 20 minutes by bus and 10 minutes by automobile,
4. In-vehicle travel times are 30 minutes by bus and 15 minutes by automobile, and
5. Out-of-pocket costs are 50 cents by both bus and automobile.

Table 1. Value of travel time in dollars per hour.

Variable	P(m, d)	P(m | d)	P(d | m)
Out-of-vehicle travel time	3.02 (1.44)	1.36 (0.98)	4.67 (4.36)
In-vehicle travel time	0.78 (0.68)	0.28 (0.66)	2.21 (2.01)

Note: The figures are for a household with annual income between \$10,000 and \$12,000. Numbers in parentheses are standard errors.

Table 2. Direct elasticities of the mode choice probability.

	Bus		Automobile	
Variable	P(m, d)	P(m | d)	P(m, d)	P(m | d)
Out-of-vehicle travel time	-1.01	-0.82	-0.13	-0.10
In-vehicle travel time	-0.31	-0.26	-0.05	-0.03
Out-of-pocket cost	-0.40	-0.91	-0.13	-0.23

The most striking variation in Table 2 is in the cost elasticity. The mode choice model derived from an estimated joint probability gives cost elasticities that are about half the elasticities computed from a recursive mode choice model.

The differences among the models could be attributed to specification errors, which affect a mode choice model and a destination choice model differently. The effects could be in the opposite directions, and therefore the joint probability model gave estimates that are in some way between the estimates of the two other models.

The marginal probabilities of the recursive models, which were formulated with composite variables, also demonstrated significant differences from the corresponding probabilities derived from the simultaneous models.

Thus, the chosen structure can make a big difference in terms of the values of the estimated coefficients. Inasmuch as there are a priori reasons to assume a simultaneous rather than a recursive structure, we should estimate the joint probabilities directly. Then, if necessary, we can derive any conditional probability.

CONCLUSIONS

Models based on disaggregate data and choice theory were estimated in the past either for a single travel choice, primarily mode choice, or for several choices but in a recursive structure. The empirical study that was conducted in this research demonstrated the estimation of a disaggregate simultaneous model. The results from the estimation of a simultaneous destination and mode choice model indicate that this approach is feasible within reasonable computation cost. Moreover, the estimation results of models with recursive structures for the same two choices show that important coefficient estimates vary considerably with the different model structures.

This empirical study was limited in scale, and it is recommended that the evidence should be extended to include alternative data sets, different trip purpose categories, a complete set of travel choices, and a more extensive set of explanatory variables (in particular, attraction description).

The empirical evidence taken together with the theoretical assumptions of a simultaneous structure and the advantages of disaggregate models suggests that future efforts in travel demand modeling should be in the direction of simultaneous disaggregate probabilistic models. Given the joint probability (from the simultaneous model), one can derive conditional probabilities and use the model for forecasting in sequential stages, corresponding with the UTMS procedure.

One of the important problems in using disaggregate models for forecasting is the aggregation problem. Future research efforts should investigate this problem. However, for the short run, simplified aggregation procedures, such as market segmentation, are available and can be used.

The use of disaggregate models suggests new emphasis in data collection efforts for transportation planning. The amount of data needed for disaggregate models has not yet been determined, but it is clear that a change in the general method of collecting travel data is appropriate. The comprehensive home interview survey covering an entire planning region might be replaced by several more descriptive small samples, in selected areas of the region. Thus, the emphasis should be to represent the full range of socioeconomic characteristics affecting travel behavior, rather than to sample all parts of the region at a uniform rate. Smaller scale surveys will make possible the collection of the detailed information (not conventionally collected) important for disaggregate demand models. For example, information on car pooling, how often a trip is made (instead of reporting only the trips made during the last 24 hours), institutional constraints such as preferred arrival time, and so forth, would be obtained. In addition to the travel data requirements, better information is also needed with respect to the attributes of alternative trips. In particular, the attraction data available from conventional data sources used in urban transportation planning are not very descriptive. More detailed attraction data are needed to achieve better predictions of destination choice.

In-depth studies of travel behavior based on detailed interviews and attitudinal data could be fruitful. However, it appears that the most beneficial directions for research toward improvements of transportation planning capabilities are the aggregation problem, behavioral modeling of round trips with non-home-based links, and experimental application of simultaneous disaggregate models to case studies of important transportation issues at different levels of planning.

In conclusion, this research has indicated the desirability and the feasibility of a simultaneous disaggregate travel choice model.

ACKNOWLEDGMENTS

During the course of this research I have benefited from the assistance and contributions of many individuals. Marvin Manheim has contributed significantly to the development of this study. His constant encouragement and keen insights were invaluable. Richard de Neufville and Wayne Pecknold provided many valuable ideas and suggestions. The assistance and advice received from Robert McGillivray, of the Urban Institute, also contributed greatly to the development and the presentation of the ideas in this research.

I would also like to thank my colleagues and friends at M.I.T., in particular, Earl Ruiter, Frank Koppelman, and Len Sherman for their constant help along the way. Daniel Brand, Gerald Kraft, and Thomas Domencich also contributed many ideas through their previous work in the subject area.

This research was supported in part by a grant to M.I.T. from the Ministry of Transport of the State of Israel for transportation planning research. Early stages of the research were supported in part by a Sloan Research Traineeship from M.I.T.

REFERENCES

1. Aldana, E. Towards Microanalytic Models of Urban Transportation Demand. M.I.T., Cambridge, PhD thesis, 1971.
2. Ben-Akiva, M. E. Structure of Passenger Travel Demand Models. M.I.T., Cambridge, PhD thesis, 1973.
3. Brand, D. Travel Demand Forecasting: Some Foundations and a Review. HRB Special Report 143, 1973.
4. Charles River Associates, Inc. A Disaggregate Behavioral Model of Urban Travel Demand. Federal Highway Administration, U.S. Department of Transportation, 1972.
5. Domencich, T. A., Kraft, G., and Valette, J. P. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968.
6. Urban Transportation Planning, General Information and Introduction to System 360. Federal Highway Administration, U.S. Department of Transportation, 1970.
7. Fleet, C. R., and Robertson, S. R. Trip Generation in the Transportation Planning Process. Highway Research Record 240, 1968.
8. Kraft, G. In Demand for Intercity Passenger Travel in the Washington-Boston Corridor. U.S. Department of Commerce, 1963.
9. Kraft, G., and Wohl, M. New Directions for Passenger Demand Analysis and Forecasting. Transportation Research, Vol. 1, No. 3, Nov. 1967.
10. Luce, R. D., and Suppes, P. Preference, Utility and Subjective Probability. In Handbook of Mathematical Psychology (Luce, R. D., Bush, R. R., and Galanter, E., eds.), Vol. 3, John Wiley, 1965.
11. Malinvaud, E. Statistical Methods of Econometrics. Rand McNally, 1966.
12. Manheim, M. L. Practical Implications of Some Fundamental Properties of Travel-Demand Models. Highway Research Record 422, 1973.
13. Martin, B. V., Memmott, F. W., and Bone, A. J. Principles and Techniques for Predicting Future Urban Area Transportation. M.I.T. Press, Cambridge, 1961.
14. McCarthy, G. M. Multiple Regression Analysis of Household Trip Generation—A Critique. Highway Research Record 297, 1969.
15. McFadden, D. The Revealed Preferences of a Government Bureaucracy. Institute of International Studies, University of California, Berkeley, Technical Report 17, 1968.

16. Plourde, R. P. Consumer Preference and the Abstract Mode Model: Boston Metropolitan Area. Department of Civil Engineering, M.I.T., Cambridge, Research Report R68-51, 1968.
17. Quandt, R. E., and Baumol, W. J. Abstract Mode Model: Theory and Measurement. Journal of Regional Science, Vol. 6, No. 2, 1966.
18. Reichman, S., and Stopher, P. R. Disaggregate Stochastic Models of Travel-Mode Choice. Highway Research Record 369, 1971.
19. Talvitie, A. Comparison of Probabilistic Modal-Choice Models: Estimation Methods and System Inputs. Highway Research Record 392, 1972.
20. Weiner, E. Modal Split Revisited. Traffic Quarterly, January 1969.

DISAGGREGATE ACCESS MODE AND STATION CHOICE MODELS FOR RAIL TRIPS

Peter S. Liou, New York State Department of Transportation; and
Antti P. Talvitie, University of Oklahoma

In this study disaggregate probability choice models are developed for access mode and for access station selection. In each of the models, there are at least two alternatives available to the individual traveler. A multinomial logit model that is based on the axiom of the "independence of irrelevant alternatives" is used. Two methods of approach concerning travelers' decision-making processes are used. The first is the simultaneous approach, which assumes that the traveler may make the access mode and station choice decisions in one of two sequences: station choice preceding mode choice or mode choice preceding station choice. In the sequential approach, the choices of access mode and access station are modeled separately. Results suggest that the traveler's decision-making process for the access mode and station choices is behaviorally separate, the sequence being station choice followed by access mode choice. The study also shows that travelers do not assign the same weights to the set of transportation system attributes when making these decisions and that the pedestrian and bus modes are preferred to the automobile mode. For the station choice, the accessibility of the train station has the greatest effect on the traveler's decision.

•A PERSON planning any type of an intraurban trip makes a number of choices including those on destination, mode, and travel route. These decisions have an important bearing on transportation planning, and therefore the knowledge of how travelers go about making their decisions is essential to transportation planners.

This research discusses the access part of the rail journey. It is assumed of course that decisions on trip origin, trip destination, rail line, and so on have already been made; consequently, travelers are faced with two access choices: access mode and access station.

The main purpose of this study is to develop disaggregate choice models of the access mode and station selection for rail work trips. At the same time, this study investigates whether travelers make the two choices simultaneously or in a sequence and, if the latter, which specific sequence? Another objective of this study is to determine the types of transportation and socioeconomic attributes that affect travelers' choice decisions and how much.

FORMULATION OF ACCESS MODE AND STATION CHOICE MODELS

In this study, a multinomial logit model (15) is used because there are usually at least two alternatives available in each choice. The generalized expression of the multilogit model is

$$P_i = \frac{e^{U_i}}{\sum_j e^{U_j}} \tag{10}$$

where

P_i = probability of an individual choosing alternative $i \Sigma \{j\}$ and
U_i = utility function associated with alternative i.

The utility of an alternative travel choice is represented by attributes of the alternative (e.g., travel time, travel cost) and the socioeconomic attributes of the individual (e.g., income). A basic assumption to formulation of these models is that travelers rationalize their choices by selecting alternatives with the highest utilities.

Construction of the access mode and station choice models in this study is approached from two behavioral process assumptions: (a) the simultaneous assumption in which the choices of access mode and station are made together and in which the joint probability of the two choices, $P(m, s)$, is contained in a single model and (b) the sequential assumption in which station and access mode choices are made one at a time. In the latter case, the joint probability of the two choices is the product of a conditional probability of one choice and a marginal probability of the other choice, e.g., $P(m, s) = P(m | s) \times P(s)$ for the station-mode sequence and $P(m, s) = P(s | m) \times P(m)$ for mode-station sequence. Clearly, both sequences can be justifiably assumed, and therefore both are studied.

For the station-mode sequence, the (conditional) mode choice is modeled first, where utility is a function of the level-of-service attributes of the mode and mode-specific socioeconomic attributes of the traveler. The (marginal) station model, on the other hand, contains three types of station level-of-service attributes: station accessibility level of service; in-train level-of-service difference resulting from choosing the selected station instead of other stations in the vicinity of the trip origin; and intrinsic level of service of the station, such as parking facility. The last two types of attributes may be directly obtained. However, accessibility to a station is related to the effort that is required of the individual to reach the station by the available modes. Therefore, it is obtained by combining the probabilities of choosing the access modes with either the access mode level-of-service variables or the utility associated with each access mode. The former method results in a number of weighted modal level-of-service variables called weighted price variables. The latter method produces a single variable called the weighted inclusive price variable. In the weighted price method the same attributes may have different coefficients in the estimated models at the two choice levels, whereas in the weighted inclusive price method the values of the coefficients for the modal level-of-service variables remain unchanged. This is because the entire utility function in the access mode model is weighted in the weighted inclusive price method. From a behavioral standpoint, the weighted price method simply assumes that travelers value the same modal level-of-service attributes differently when they choose modes and stations. On the other hand, the weighted inclusive price method assumes that travelers view the relative importance of the modal level-of-service attributes equally at the two choice levels. An interesting consequence of this method of approach, as reflected in the estimated weighted inclusive price station model, is that, if this assumption is valid, then the coefficient of the weighted inclusive price variable should be 1.

The other possible sequence is the mode-station sequence in which the conditional station choice $P(s | m)$ is modeled first and then the marginal mode choice $P(m)$ is modeled.

In the simultaneous model structure, the probability that a traveler chooses access mode m and access station s is a function of the level-of-service variables of each available mode, the in-train level-of-service difference variables, and the intrinsic station variables. Again, the socioeconomic attributes of the traveler are included in the level-of-service variables.

DATA AND METHODS

Data Source and Sample Selection

The trip data used in this study were taken from an origin-destination survey conducted for the Southward Transit Area Coordination (STAC) committee in Chicago. The inner part of the STAC area (21) was chosen as the study area because most rail work trips in the area originate there. The trip origins may be identified from the survey by a ¼-square mile centroid. The access mode, access station, and the access distance to the station can also be either directly or indirectly obtained from the survey.

A total of 150 work trips were randomly selected from the Illinois Central (IC) Railroad surveys. Those samples with incomplete information were replaced with valid samples, also randomly picked. Sets of 25 samples each were selected in a similar manner from the Rock Island (RI) and the South Shore and South Bend (SS) Railroad data. Table 3 gives the number of travelers by access mode and rail line.

Construction of the access mode and station selection models is based on data on travelers' work trips on the Illinois Central. Rock Island and South Shore data are used only to test the various operational models.

The dependent variable of a multinomial logit model is the choice probability, P_i, where i is one of the alternatives in the choice set. Because only the actual choice is observed and not the probabilities, when model parameters are estimated P_i equals one for the chosen alternative and zero otherwise.

For the access mode choice models, the alternatives considered are automobile, bus, and walk. (Even though it was determined whether travelers drove or were driven, the data did not permit further detail in modeling access mode choices.) However, each person in the sample was not always considered as having all alternatives available to him. Automobile (driven or drove) was always considered a relevant alternative. Walking was considered to be unavailable to a person if walking distance to a station was more than 20 minutes. The bus mode was available if a traveler was within a ½-mile walk of a bus route.

For the station selection model, alternative stations were chosen on the basis of data and were usually near the chosen station.

Notation

The notation (22) used in the models is defined as follows:

OVT = out-of-vehicle time, the sum of walking time and waiting time during the individual's access trip to the station;
AT = automobile time, the amount of time the individual spends in an automobile during his station access trip;
BT = bus time, the amount of time the individual spends on a bus during his station access trip;
OC = operating cost of an automobile during the access trip to the station;
PC = out-of-pocket parking cost for automobile driver or bus fare for the bus user;
TC = total cost, the sum of the operating cost and the out-of-pocket cost;
LHT = line-haul time difference, the on-train travel time difference resulting from choosing the station instead of the alternative stations;
PD = parking dummy of the available parking space per automobile driver; and
S = socioeconomic attribute, the ratio of total cost to median income.

Table 3. Sample distribution.

	Railroad		
Access Mode	IC	RI	SS
Automobile			
Driver	27	4	4
Passenger	20	4	4
Walk	50	9	9
Bus	53	8	8
Total	150	25	25

Estimation Method and Evaluation Criteria

The models are evaluated in three ways: (a) The statistical significance of each variable in the model and the model as a whole is determined; (b) the reasonableness of the magnitude of the coefficients of the model variables is examined (elasticities are studied to determine the effects of the attributes on the choice probability and the value of time); and (c) the model is applied to situations different from that on which the model is estimated.

Inasmuch as the choice probabilities are not observed, a statistical test such as the estimated residue measurement (R^2) ordinarily used for linear regression analysis is not valid. The statistical tests used for the disaggregate models in this study are mainly the t-test, which determines the statistical significance of each variable in a model, and the χ^2-test, which determines the statistical significance of the entire model.

Both the sign and the magnitude of the coefficient of a variable are examined. The sign of a coefficient must be logical: The coefficients for out-of-vehicle time, automobile time, bus time, operating cost, out-of-pocket cost, weighted price, and socioeconomic variables must have negative signs, whereas the coefficients of line-haul difference, parking dummy, and weighted inclusive price variables must have positive signs.

The magnitude of the coefficients of the variables may be examined by studying the elasticities of the choice probabilities with respect to each of the variables. The mathematical expression for the direct and cross elasticities of a multilogit model are

$$E_{X_{li}} = b_l X_{li}(1 - P_i) \tag{11}$$

$$E_{X_{lj}} = -b_l X_{lj}(P_j) \tag{12}$$

where

P_i = choice probability of alternative i,
X_{li} and X_{lj} = lth explanatory variable describing alternatives i and j,
b_l = coefficient of X_l, and
$E_{X_{li}}, E_{X_{lj}}$ = direct and cross elasticities with respect to X_l.

Furthermore, the implied value of time obtained from this research is compared with the value of time obtained from other studies.

The disaggregate access mode and station models are further evaluated by applying each model to different situations. As mentioned previously, the base data for the models estimated in this study are the set of IC data; the RI and SS data are the control data and are used solely for testing the models. The service areas, operators, number of rail tracks, distances between adjacent stations, train operating frequencies, and types of signal and train facilities of the Rock Island and South Shore Railroads are different from those of the Illinois Central Railroad.

For each individual sample, the expected probability of the chosen mode or station is compared with the expected probabilities of the alternative modes or stations in their respective choice set. If the expected probability of the chosen mode or station is greater than or equal to those of the alternatives, then the model has made a correct prediction. Otherwise, the prediction is wrong. Furthermore, for the access mode model, the expected number of users of each mode is compared with the actual number of users of the same mode.

RESULTS AND EVALUATION OF ESTIMATED ACCESS TRIP CHOICE MODELS

Models were estimated for the conditional mode choice P(m | s), the marginal station choice P(s), and the conditional station choice P(s | m). However, estimation of the marginal mode choice model P(m) and the simultaneous choice model P(s, m) resulted in models with incorrect signs. These models and evaluations of them are discussed below.

Conditional Mode Choice Model

Two of the estimated access mode models appeared to have the correct signs and statistically acceptable coefficients for each variable. The coefficients of these two models and other relevant information are given in Table 4. Both of these models include the out-of-vehicle, automobile, and bus times. The cost variable is different, however, in the two models. In the first model it is the operating cost (OC), and in the second model it is the total cost divided by income (S).

Statistical tests indicate that all the variables in model 1 and the model itself are significant at the 0.99 level of confidence. In model 2, the socioeconomic variable is statistically significant only at the 0.75 level of confidence. The bus time variable is statistically significant at the 0.95 level of confidence. The out-of-vehicle and automobile time variables, along with the model itself, are statistically significant at the 0.99 level of confidence. Therefore, on the whole both models are statistically acceptable.

From the coefficients in model 1, the implied values of time (in dollars per hour) are as follows:

Variable	Value	Standard Deviation
Out-of-vehicle time	0.48	0.25
Automobile time	0.75	0.50
Bus time	0.41	0.25

Comparisons of these values with the submode values of time in other studies are not available. However, the value of in-vehicle time is approximately the same in this research and in some other recent demand model studies, approximately 70 cents/hour (2, 19). However, the values of the out-of-vehicle time in this research are much lower than the value of the out-of-vehicle time of other studies, $3.00/hour and more. It should be noted, however, that the trips under investigation in this study are access trips, whereas the other studies considered either the major part of the trip or the entire trip.

From the coefficients of the second model, the implied value of automobile time is $80/hour, which is too large to be reasonable. Therefore, model 2 is considered invalid.

The direct elasticities of the access mode model, computed at the means of each variable, for each fixed probability are given in Table 5. It can be seen from Table 5 that most of the variables are elastic when computed at the means of these variables. OC, which is associated exclusively with the automobile mode, has the greatest elasticity, and AT, OVT, and BT have smaller elasticities in that order. Values of the direct elasticities in this study are not in line with the a priori knowledge of the elasticities from previous studies. However, the differences between this study and others must again be noted. Also, as the probability increases, the elasticities decrease. This suggests logically that travelers grow less concerned with changes in the transportation attributes if their chosen mode is chosen with a high probability. The model is further evaluated by applying it to both IC and RI/SS data and comparing the results. The misclassifications and the predictive rates are given in Table 6.

The expected number of travelers by mode can be computed as the sum of the expected probability values of each mode:

$$N_m(\text{expected}) = \sum_i P_m^i$$

where

P_m^i = probability of mode m being chosen by person i and
$N_m(\text{expected})$ = expected number of travelers to use mode m.

Comparisons of the expected and the actual number of travelers by mode are given in Table 7.

The comparisons show that the expected and actual values by mode are compatible. For the RI and SS data, the absolute percentage of difference for bus mode is 69 as compared to approximately 80 for the other modes. This may be attributed to the fact that bus frequencies in the areas of the RI and SS Railroads are often quite low. Though the waiting time for bus was set at no more than 8 minutes during the process of data preparation, 30-minute headways for buses in these areas are not uncommon. This may interfere with a traveler's time schedule for reaching the station and eventually the jobsite and therefore force him to choose another access mode.

Marginal Station Choice Model

Two of the estimated station models appeared to have correct coefficient signs and statistically acceptable indications. One of the models used the weighted prices and the other used the weighted inclusive price as part of their level-of-service variables.

The Weighted Price Station Model—The statistical tests of this model (Table 8) indicate that the variables are significant at the following levels of confidence:

Variable	Level of Confidence
Weighted OVT	0.99
Weighted AT	0.99
PD	0.95
LHT	0.80
Whole model	0.99

The direct elasticities are computed at the means for the weighted OVT and weighted AT, at 4 minutes for LHT, and at 1 minute for PD (Table 9).

These elasticities indicate that, when selecting the access stations, travelers are most sensitive to out-of-vehicle time and automobile time. The results also show that travelers are relatively unconcerned about the extra amount of time spent (or saved) inside the train in choosing the access station. In spite of the incompleteness of the parking availability variable, it appears that it has an effect on the choice of access station. Information on the value of time is not available, inasmuch as this model has no cost variable. Misclassifications and the predictive accuracy rates of the model are given in Table 10.

The Weighted Inclusive Price Station Model—The coefficients of the weighted inclusive price station model and other relevant information are given in Table 11. Statistical tests of this model indicate that the variables are significant approximately at the levels of confidence given below:

Variable	Level of Confidence
Weighted inclusive price	0.97
PD	0.97
LHT	0.90
Whole model	0.99

The direct elasticities are obtained in the same way as in the previous station model (Table 12). The elasticities indicate that the weighted inclusive price variable is the most important attribute to travelers selecting a station.

The coefficient of the weighted inclusive price variable in this model is 0.5850. It is tested to be significantly different from 1.0000. This indicates that the above-mentioned assumption is invalid. In other words, the traveler does assign different weights to the set of transportation system attributes when making his access mode and station choice decisions.

The misclassifications and the predictive accuracy rates of the model are given in Table 13. Comparisons of the actual number of travelers choosing a certain train station with the expected number are not made for the two access station models because the small number of travelers observed is distributed to a relatively large number of alternative stations.

Table 4. Coefficients of conditional mode models.

Variable	Model 1[a] Coefficient	Model 1[a] Standard Error	Model 2[b] Coefficient	Model 2[b] Standard Error
Out-of-vehicle time	-0.441	0.094	-0.286	0.059
Automobile time	-0.681	0.291	-1.122	0.260
Bus time	-0.382	0.120	-0.164	0.091
Operating cost	-0.556	0.193	—	—
Socioeconomic attribute	—	—	-0.084	0.111

[a]χ^2 = 99.498 with 4 degrees of freedom.
[b]χ^2 = 87.647 with 4 degrees of freedom.

Table 5. Direct elasticities of conditional mode model.

Variable	Mean (minutes or cents)	Direct Elasticity P = 0.30	P = 0.50	P = 0.70
OVT	6.7	2.08	1.49	0.89
AT	6.7	3.19	2.28	1.37
BT	10.1	2.71	1.94	1.16
OC	15.5	6.03	4.31	2.58

Table 6. Accuracy of conditional mode model.

Data	N	M	$\alpha = 1 - \frac{M}{N}$ (percent)
IC	150	12	92
RI/SS	50	7	86

Note: N = total number of observations, M = number of misclassifications, and α = predictive accuracy.

Table 7. Comparison of number of mode users.

Data	Mode	N_E	N_A	Absolute Percentage of Difference
IC	Automobile	56	47	81
	Walk	40	50	80
	Bus	54	53	98
RI/SS	Automobile	13	16	81
	Walk	15	18	83
	Bus	21	16	69

Note: N_E = expected number of travelers and N_A = actual number of travelers.

Table 8. Coefficients of weighted price station model.

Variable	Coefficient	Standard Error
Weighted OVT	-0.385	0.102
Weighted AT	-0.957	0.172
LHT	0.138	0.167
PD	0.827	0.469

Note: χ^2 = 90.391 with 4 degrees of freedom.

Table 9. Direct elasticity of weighted price station model.

Variable	Variable Value (minutes)	Direct Elasticity P = 0.30	P = 0.50	P = 0.70
Weighted OVT	6.44	1.73	1.24	0.74
Weighted AT	2.16	1.44	1.03	0.62
LHT	4.00	0.39	0.28	0.17
PD	1	0.57	0.41	0.25

Table 10. Results of application of weighted price station model.

Data	N	M	$\alpha = 1 - \frac{M}{N}$ (percent)
IC	150	32	78.7
RI/SS	50	1	98.0

Table 11. Coefficients of weighted inclusive price station model.

Variable	Coefficient	Standard Error
Weighted inclusive price	0.585	0.107
LHT	0.230	0.194
PD	1.189	0.563

Note: χ^2 = 100.4353 with 3 degrees of freedom.

Table 12. Direct elasticities of weighted inclusive price station model.

Variable	Variable Value	Direct Elasticity P = 0.30	P = 0.50	P = 0.70
Weighted inclusive price	8.07	3.31	2.36	1.42
LHT	4.00	0.64	0.46	0.28
PD	1	0.83	0.59	0.36

Table 13. Results of application of weighted inclusive price station model.

Data	N	M	$\alpha = 1 - \frac{M}{N}$ (percent)
IC	150	27	82
RI/SS	50	2	96

Table 14. Coefficients of conditional station models.

Variable	Model 1 Coefficient	Model 1 Standard Error	Model 2 Coefficient	Model 2 Standard Error
OVT	-0.664	0.180	-0.653	0.177
BT	-0.310	0.208	—	—
In-vehicle time	—	—	-0.184	0.173
OC	-0.136	0.066	-0.062	0.094
PD	0.859	0.480	0.833	0.480

Table 15. Accuracy of conditional station models.

Model	Data	N	M	$\alpha = 1 - \frac{M}{N}$ (percent)
1	IC	110	20	82
	RI/SS	23	3	87
2	IC	110	20	82
	RI/SS	23	3	87

Table 16. Marginal mode choice models.

Variable	Model 1	Model 2	Model 3
OVT	-0.083	-0.082	-0.347
BT	0.137[a]	—	—
In-vehicle time	—	0.142[a]	0.210[a]
OC	-0.189	-0.248	-0.093
D1	—	—	2.837
D2	—	—	3.638

Note: D1 and D2 are the dummy variables such that D1 = 0, D2 = 0 for automobile mode, D1 = 0, D2 = 1 for walk mode, and D1 = 1, D2 = 0 for bus mode.
[a]Incorrect coefficient sign.

Conditional Station Selection Model

Estimation of the conditional station model is the first step of the mode-station modeling sequence. The coefficients of the two models and other relevant statistical information are given in Table 14.

The values of time implied by the coefficients of model 1 are approximately \$3.00 at $\sigma = 1.73$ for OVT and \$1.40 at $\sigma = 1.12$ for BT. The values of time implied by the coefficients of model 2 are approximately \$6.00 at $\sigma = 9.94$ for OVT and \$1.80 at $\sigma = 4.20$ for in-vehicle time. The misclassifications of these two models when applied to the base and the control data are given in Table 15.

These two station models appear to be fairly good. Nevertheless, it must be noted that the station models only constitute part of the sequential modeling process. The access mode models also have to be examined before the validity of this particular mode and station decision-making sequence assumption can be determined.

Marginal Mode Choice Model

The level-of-service variables describing access to a station by mode were aggregated by the weighted price method. All the models involved incorrect coefficient signs (Table 16). The fact that no valid access mode model could be estimated raises doubt about the validity of the mode-station decision-making sequence assumption. Therefore, the marginal mode models as well as the conditional station models cannot be applied with confidence to planning problems.

Simultaneous Choice Model

No valid models could be obtained by using the simultaneous model structure. An example of the estimated simultaneous models is given below:

Variable	Coefficient
OVT	-0.0119
AT	-0.0571
BT	0.1974
PD	1.340

(Note that the coefficient for bus time shows an incorrect sign.) The results tend to suggest that the simultaneous model structure is also an invalid traveler decision-making assumption.

SUMMARY

The main purpose of this study was to develop disaggregate choice models of the access mode and access station for those travelers who make their work trip by rail.

The multinomial logit model, which is used in this study, is based on the independence of irrelevant alternatives assumption and is capable of dealing with a different number of choice alternatives for each of the behavioral units; it is considered to be the most suitable model for the situation under investigation.

The data used for the estimation and evaluation of the various probability models were obtained from the Chicago area.

It is assumed that a person makes the access mode and station choice decisions either simultaneously or in the station-mode sequence or the mode-station sequence. In the case of the sequential assumption, the joint probability of the access mode and station is separated into a conditional probability of one choice given the other choice and a marginal probability of the other choice, depending on the particular choice sequence assumed. The investigation in this study of the simultaneous model structure and the mode-station sequence structure failed to produce choice models with intuitively correct coefficient signs.

The results of the research based on the station-mode sequence revealed some interesting behavioral characteristics of individual travelers when they make their access trips. Even though some studies have reported rather high cost elasticities for the automobile mode and rather low elasticities for out-of-vehicle and in-vehicle time (2, 19), it is a common belief that travel demand is insensitive to changes in travel cost and possibly in-vehicle time, although it is quite sensitive to changes in out-of-vehicle time (2, 5, 23). It should be noted, however, that the latter are derived from travel demand models, whereas the former are so-called modal-split elasticities (i.e., trip frequency decisions were not modeled).

The results from the access mode model, $P(m|s)$, of the present study indicate that, of the travel time (modal-split) elasticities, the automobile time elasticity is the highest followed by bus time and out-of-vehicle time elasticities. Surprisingly, the automobile operating cost elasticity is the highest of all by a wide margin, and several attempts to include automobile ownership costs and bus fare in the models failed to produce plausible models. Finally, the value of automobile time was estimated at 74 cents/hour; this is in accordance with previously obtained values.

An income variable was also considered in the estimation of one of the access mode models. A very rough income figure, the median income of the traffic zone, was the only available income information. It is not known whether this is why the model in which this variable was included yielded an implausible model; income information for each individual would have been desirable.

These results indicate that paying for a car trip to a station and spending travel time inside an automobile are disliked by travelers as compared to spending travel time inside a bus. The relatively low elasticity of out-of-vehicle time (as compared to automobile operating cost) suggests, furthermore, that it should not be difficult to convince a traveler to choose access modes such as walking and even the bus for which the out-of-vehicle time may constitute a large part of the total price.

These "discoveries" appear contradictory to the information obtained from previous travel demand studies. Nevertheless, it must be realized that those studies considered the entire trip, whereas this study deals only with the access part of the trip. These two are different in nature, and consequently the behavioral responses of the travelers should not be expected to be the same.

An important consideration in this context is whether access trips can be separated from the rest of the journey. This assumption was made in this study, but it is by no means the only assumption that can be made. In similar vein, whether automobile ownership and location decisions of households should be linked with the work trip decisions must also be asked. This research does not provide answers to such questions, of course, but the somewhat counter-intuitive results tend to suggest that automobile ownership and location decisions are important in work trip decisions (and vice versa).

Still, the different behaviors may be intuitively justified. For the entire journey, the travel distance between the trip origin (home) and the trip destination (jobsite) is generally quite large. Therefore, comfort, privacy, and other advantages offered by the automobile mode become important to travelers and thus make their response insensitive to changes in automobile travel time and cost characteristics. Of course, the walk mode is usually not considered as one of the available alternatives for such a trip. On the other hand, for the access trip, the travel distance between the trip origin and the trip destination (train station) is very short. For example the average access trip travel distance of the 150 observed travelers in the set of base data is only 1.5 miles. Clearly, not much comfort or privacy can be derived by using a car for a trip of this length. On the contrary, the various inconveniences of using an automobile, such as finding a parking space, leaving the car in a parking lot close to home where it is not available for use by other members of the family, or having someone else drive the traveler to the station, become predominant disadvantages.

Access station selection models, P(s), were more in keeping with current beliefs about travel behaviors: Travelers' choice decisions are most sensitive to out-of-vehicle time followed by the automobile time, whereas bus time and travel cost variables failed to enter the model. Also, in the weighted inclusive price station selection model, the coefficient of the inclusive price variable is significantly different from 1.0, which suggests that travelers do not assign the same weights to the set of transportation system attributes when they choose access mode and station.

In regard to the simultaneous and mode-station sequential models, no conclusive explanation can be given of their failure to obtain plausible choice models. The results of this research only give empirical support to such travelers' decision processes in which the access mode and station choice is done sequentially—station choice followed by mode choice. This is, of course, a tentative suggestion.

Finally, when the access mode and station selection models were applied to different situations they produced good predictive results.

WHAT HAS BEEN LEARNED

It has been learned through this research that more detailed information than was available on the level of service of the transportation system and on the individuals in the sample is required in order to estimate disaggregate choice models effectively. Within the extent of this study, for example, the exact location of the trip origin, the individual's income, specific information on parking conditions at the stations, and most importantly relevant alternatives to the choices actually considered by each individual behavioral unit should be specifically determined when surveys are taken for disaggregate choice modeling. Of specific concern is the automobile mode, which in this study is considered a relevant alternative for every traveler.

The somewhat counter-intuitive results also suggest that there is a clear need to relate household location and automobile ownership decisions to choice of (access) mode and other (work) trip decisions if truly behavioral models are to result.

In conclusion, the disaggregate modeling technique and the information obtained were quite instructive. Only a small sample of data was required to estimate the models; thus, considerable savings of money and time can result from the use of disaggregate models. However, before disaggregate models can be confidently used in transportation planning, their transformation into aggregate travel demand models must be accomplished. To date, little work (24) has been done on forecasting aggregate travel demand by means of (transformed) disaggregate models. Of the few aggregation procedures, more empirical studies are warranted.

ACKNOWLEDGMENTS

This research was partially supported by the Urban Mass Transportation Administration, U.S. Department of Transportation. Appreciation is extended to Moshe Ben-Akiva of M.I.T. for supplying the computer program necessary for this work. The authors retain responsibility for any errors and omissions in the paper.

REFERENCES

21. Southward Transit Area Coordination Study Report. Chicago, 1970.
22. Liou, P. S. Disaggregate Access Mode and Station Selection Models for Rail Trips. Department of Civil Engineering, University of Oklahoma, PhD dissertation, 1973.
23. Talvitie, A. P. A Direct Demand Model for Downtown Work Trips. Transportation, No. 2, 1973.
24. Talvitie, A. P. Aggregate Travel Demand Analysis With Disaggregate or Aggregate Travel Demand Models. Proc., Transportation Research Forum, 1973.

DISCUSSION

Gerald Kraft, Charles River Associates, Cambridge, Massachusetts

I want to clarify an issue raised in both papers that leads to serious confusion. Both papers contrast simultaneous models with what they call sequential models. Unfortunately, the urban transportation planning models of the generation-distribution-modal split-route assignment variety give particular meaning to the issue of simultaneity and sequentiality that is not the same as those presented in these papers. Whereas the papers address a sequential decision-making process, the mathematical formulations used imply no such concept. These formulations are simply a factoring of a joint probability distribution into the product of a conditional and a marginal distribution. Thus, it is a process of estimating probability functions sequentially and does not imply anything about decision-making in any sequence.

Recursive, the term used by Ben-Akiva, is far more appropriate. Although it is of course true that travelers may make sequential decisions, there is no particular reason to assume that they do. In this sense Ben-Akiva, arguing for simultaneous structures, is quite correct. On the other hand, to assume a joint probability distribution and then ask questions about conditional probabilities or marginal probabilities are perfectly natural and reasonable. In any event, I would prefer using recursive to sequential to avoid any possible ambiguity.

GENERAL REMARKS

Both papers are very interesting and appear to provide us with some very useful information. A puzzle appears when they are regarded together. In effect, Ben-Akiva says that the recursive approach to modeling does not work because there is no a priori reason to have the model go in one order, say destination to modal split, rather than the other, modal split to destination. He shows that, by changing the order, he obtains disparate empirical results. He then argues that a simultaneous estimation not only avoids the ambiguity of ordering the conditional distributions to be estimated but also is mathematically and statistically feasible for estimation.

Liou and Talvitie also look at two alternative recursive structures, one in which access mode precedes station selection and the other in the opposite order. They also find the results different, as does Ben-Akiva, but, when they attempted a simultaneous model, no useful results were obtained.

Thus, although both papers agree that the two recursive structures give different answers, one paper argues that the solution is to use simultaneous estimation and the other says that simultaneous modeling produced no useful results. Furthermore, Liou and Talvitie argue that the order is important and care should be given to the selection of the order. They even suggest that the order can be determined from the empirical results.

Neither paper really addresses the question of why the authors came out with the results they did. Neither Ben-Akiva nor Liou and Talvitie tell us why the different orderings on the recursion yield substantially different results, nor do they tell us why the results obtained by using recursive estimation differ from those that were obtained by using simultaneous equations.

Before going into the possibilities, I should state a basic inconsistency in the Liou-Talvitie paper that may be the root of at least part of the problem. They assume that the station and access mode decision process can be divorced from the selection of the basic line-haul mode. In a sense, by assuming such independence they have already assumed the appropriateness of the recursive system and may have thus "cooked" the results. This may explain in part why their simultaneous model does not work. The authors recognize this, but perhaps they do not place enough emphasis on the problem in their interpretation of the results using the recursive model.

In evaluating the reasons for discrepancies between the recursive and simultaneous approaches, I see four possibilities.

1. The theory and assumptions used in constructing the model itself may be inadequate.
2. The specification of a structure may create problems.
3. The variables used in the model may be inadequate or inappropriate.
4. The techniques used for estimating the parameters may be inadequate.

Before I begin to consider these possibilities in turn, I apologize for not having a specific answer to the confusion, but I hope my discussion can be useful as a guide to seeking the reason.

MODEL THEORY

In considering the problems of theory, we must examine probability theory and utility theory. In probability theory, the equivalence between simultaneous and recursive estimates as defined by the authors appears simple; as I indicated earlier a joint probability distribution can be expressed as the product of a conditional distribution and a marginal distribution. Furthermore, in a multiple decision framework, repeated application of this process can be used to develop a chain of probabilities that can be modeled.

In another but similar context, Manheim demonstrated the equivalence of simultaneous and sequential or recursive models (12). As for probability theory at this level, the joint probability function is independent of the order of recursion. Thus the results obtained in the papers are totally inconsistent with the probability theory, which I believe we all would accept as given.

On the other hand, the utility theory on which the models are based might be questioned. This concerns principally the assumptions of separability or additivity in the utility functions. (These assumptions are too technical to take up here, but they are extremely powerful tools for simplifying the development of models and for enhancing the power of their application.) They may, nevertheless, do violence to reality with the consequence that they lead to the kinds of results presented here. This suggests some merit in further investigation but it is not likely that this is the most fruitful first step.

MODEL STRUCTURE

Have the theoretical considerations of probability and utility been applied properly in development of a specific structure for the model? The model development leading to the logit models used by the authors and in other work appears to be quite sound. As a fairly general form for an ogive it would appear to be very robust. The specific forms of the variables entering the model, however, may not be entirely appropriate. There may be interdependencies, for example, that are not taken into full account in the types of linear functions that have been explored.

Whereas the logit model has substantial advantages because of its ease in manipulation and ease with which new alternative choices are added, it may be too simplified and we may have to forgo some of its benefits. We want to be extremely careful before discarding the model, however, because the advantages seem to be so overwhelming. In addition to the ease of adding alternatives to modes, destinations, and the like when the logit model is used, we must also consider possible future developments to add choices of an intermediate-run nature such as automobile ownership or a long-run nature such as the location of residence or place of work. The axiom of independence of irrelevant alternatives may be a mixed blessing, but before throwing it out we should be sure that we are unable to make the judicious decisions about alternatives that would avoid conflict between the axiom and reality. That assumes, of course, that we can find no similarly endowed alternative.

VARIABLE SELECTION

The variable selection and definition used in the models also may lead to the results obtained. Leaving out important variables or introducing spurious ones may have serious ramifications for model misspecification. It is well known that model misspecification can lead to peculiar results; all model estimation may suffer some from this problem. One problem is the attribution of effects to the wrong variables. In the case at hand the problem may be more subtle and serious. In particular, it seems that Liou and Talvitie's failure to distinguish between automobile driver and automobile passenger trips seriously compounds the problem. Also, their use of aggregate zonal income may explain some of the inconsistencies. This is suggested, for example, by the very poor statistical results obtained for their socioeconomic variable, the ratio of total cost to median income.

The idea of a combined price or cost variable with income is an interesting one; it was also used by Ben-Akiva in his dissertation (2). The inability of such a variable to pick up separate price and income effects, however, is a serious weakness (although there may very well be a relationship between the level of income and the size of the price effect).

An additional problem may arise in the Liou and Talvitie study because no variable is provided for a difference in rail fares between the alternative stations. If there is no difference, they should tell us so, but one might suppose that such a difference (at least at the margin) would have an effect on station choice.

It seems that a fruitful avenue of exploration to explain the differences in results as we change the direction of recursion and between recursive and simultaneous estimation approaches might lie in better variable selection, which may isolate effects that are currently being related improperly.

ESTIMATION TECHNIQUES

The last area that may lead to discrepancies between recursive approaches and between the recursive approach and the simultaneous approach is in the estimation techniques themselves. The idea of an inclusive price developed by Charles River Associates (4) was to simplify the estimation process. The results obtained through the use of such a variable were extremely encouraging. In terms of parameter estimates obtained, the inclusive price was statistically quite significant, and the weight seemed to be reasonable. Ben-Akiva's dissertation (2) reaffirmed those characteristics. The use of such a function, if it is consistent with the theory, is extremely beneficial. It reduces the number of variables that must be considered at any single stage of the estimation and may make an otherwise intractable estimation problem relatively simple. Nevertheless, it may be that the process using an inclusive price creates estimation problems.

Ben-Akiva informs us that the estimation times for the recursive and simultaneous models were not very different when input-output differences were considered, the simultaneous being somewhat longer, but not significantly so. On the other hand, when the computing time itself is considered, the simultaneous model took nearly four times as long as the recursive model. With larger samples, more variables, and more stages of the recursion to be processed, such as trip frequency and time of day, the differences in computation time between methods may be substantial. Furthermore, estimating equations with large numbers of variables is an extremely difficult procedure. I am not fully convinced that we should discard recursive estimation in favor of simultaneous estimation. Rather, we should explore further the use of the inclusive price, including the possible introduction of constraints in the estimation process that will ensure consistency between the different recursion orders and between recursive and simultaneous estimation techniques.

DECIDING AMONG THE MODELS

An important question remaining is, How can we decide which of the models are good? The tests suggested by Liou and Talvitie seem to be good ones. They suggest we look at statistical significance, reasonableness, and application of the model to a new set of data to compare the estimated and observed values. I would put the reasonableness test first and insist that the results agree with theory and that parameter values have correct signs and be of reasonable magnitude with respect to a priori considerations. Statistical significance without reasonableness produces very questionable results.

Application of the three tests suggested may help us to reject one model over another; however, it does not tell us what is wrong with the rejected model nor indeed what is possibly wrong with a model that is accepted. Unfortunately, we cannot pick models in any general way by using these techniques. We can only compare model results with results in the real world. Even here, we must be extremely cautious because theory is all we have. If the results of statistical estimation are inconsistent with our theory, we may wish to explore the theory itself or, alternatively, go back to the entire estimation process to locate the difficulty. Indeed, the results reported in both papers appear to have problems with statistical significance.

USING THE MODELS

In closing, there are two problems that need a great deal more attention before we will be able to use disaggregate models generally in the planning process. Both papers indicate that further research is essential to develop these models for general application. This is not to say that we have not learned a great deal through the development and estimation of the models to date. We are making extraordinary advances in model development for urban transportation planning purposes, but it is important that we explore the causes for the problems cited in these papers.

Both authors make brief mention of the aggregation problem in the use of disaggregate models. The problems are not trivial. They cannot be dealt with by simple handwaving and suggesting that aggregation can be accomplished by estimation followed by addition of the results for many individuals. Some basic requirement is called for to examine the possibility of developing aggregate models from the disaggregate or alternatively learning how to use parameter estimates from the disaggregate models for broader, more aggregative decision-making. Indeed, this has been the genesis of these models where Lisco, Thomas, Stopher, and others attempted to measure values of time and where others have been concerned with estimates of price and service level elasticities.

The fundamental problem confronting us is, "Can we develop a model structure that does not do violence to our theories or to reality yet is both mathematically and statistically tractable?"

Peter R. Stopher, Associate Professor of Civil Engineering, Northwestern University

The impression that is gained from reading these two papers is that the authors draw diametrically opposing conclusions. Ben-Akiva claims that a simultaneous model of destination and mode choices for a shopping trip can be constructed and is to be preferred over a recursive procedure comprising two models for destination and mode choices. In contrast, Liou and Talvitie are unable to construct a satisfactory simultaneous model of station and access mode choices for the access segment of a work trip and conclude that these choices must be modeled recursively.

In my opinion, neither of these conclusions is sufficiently supported by the papers, and I must conclude that there is no basis for accepting either conclusion at this time.

These papers may be discussed from a number of viewpoints. I have elected to consider the statistical evidence supplied and will leave it to other discussants to consider matters of philosophy, structure, and the like.

First, both papers are lamentably deficient in the reporting of statistical measures of assessment and comparisons for the models. Hence, many of my comments are in the form of requests for more information. Ben-Akiva bases his conclusions on the fact that the coefficients of identical variables are numerically different in the simultaneous model from those in the recursive model. However, he does not establish whether these differences are statistically significant. Liou and Talvitie dismiss the simultaneous choice model on the grounds of an incorrect sign for one variable in each simultaneous model built, but do not establish whether the incorrectly signed coefficient is statistically significant from zero. Further testing of differences between the models is largely left alone because of the incorrectly signed coefficient. The question of statistical differences between the other coefficients remains open.

Both papers make model comparisons on the basis of derived travel time values. These values are obtained by computing the ratio of the coefficients of travel time and travel cost. Both papers report a standard deviation for these computed travel time values, but neither paper reports on the method used for computing this standard deviation. Correctly, this is determined (25) as

$$V\left(\frac{a_1}{a_2}\right) = \frac{a_2^2 V(a_1) + a_1^2 V(a_2) - 2a_1a_2 \operatorname{cov}(a_1, a_2)}{a_2^4}$$

where

a_1 and a_2 = coefficients whose ratio is being determined,
V( ) = variance of, and
cov( ) = covariance of.

If both $E(a_1)$ and $E(a_2)$ are nonzero, then this variance has a distribution. However, this distribution is likely to be seriously skewed (26) and will not permit the standard deviation to be used as a means of establishing confidence limits in the normal manner. Hence, the reporting of values of travel time and their standard deviations provides little or no information for comparison between models. Based on the actual values derived by the authors, and the untenable assumption that the ratio is normally distributed, none of the values of time reported in either paper is significantly different from zero and hence each other. Thus, I must dispute the statement by Ben-Akiva that "Estimated values of time from the simultaneous model are greater than those estimated from a mode choice model (given destination) and smaller than those estimated from a destination choice model (given mode)."

Neither of the two papers reports the values of one or more constants for the multiple-choice models. However, the estimation of one or more constants permits coefficient estimates to be made with minimal bias and also allows overall goodness-of-fit measures to be determined. In binary choice models, the constant determines the position of the logit curve in relation to the values of the fitted linear function. It serves an identical purpose (in more dimensions) for a multiple-choice logit model. In conceptual terms, the constant may be considered as providing some of the information lost by an improperly specified model (27). The lack of a constant therefore presupposes a fully specified model (i.e., a constant equal to zero) and also changes significantly the meaning of the chi-square test of model fit.

A further problem with both papers concerns the data base for the choices. In both cases, it appears that the choice sets have been assigned to the individuals in the data set and the alternative options have been provided with engineering measures of the relevant variables. Many previous studies have shown this to be behaviorally incorrect. First, it is necessary to define the perceived choice set for each individual and, second, it is the perceived attributes of the alternatives that determine behavior. Because neither of these sets of perceptions was determined, the data bases of both studies must be considered suspect. The computation of models based on relative measured attributes for an arbitrarily defined set of potential alternatives provides no behavioral information and consigns the exercises to academic esoterica.

Finally, it should be noted that, in Ben-Akiva's paper, the coefficients of the variables relating to characterization of alternative destinations are generally not significantly different from zero. This is indicated by the fact that the reported standard errors of the coefficients are generally more than half the value of the coefficients for the variables EMP_d and $DCBD_d$. However, the coefficient of each variable is a function of the variances and covariances of all variables used in the model, including those that yield nonsignificant coefficients. The presence of these variables in the simultaneous model may be the sole cause for the difference in the coefficient values from the mode choice model, where the destination variables do not appear. Similarly, the destination choice model has nonsignificant coefficients for the destination variables, which makes comparisons between the recursive and simultaneous models trivial.

In summary, neither of the conclusions drawn by these papers can be accepted unless the authors can provide much more evidence. Given the questions raised here concerning the data base and the significance of the destination descriptors in Ben-Akiva's models, it is doubtful whether the conclusions can be accepted under any circumstances. The primary contributions of the two papers are, first, to highlight the problem of model structure in disaggregate, behavioral, travel demand models and, second, to demonstrate that it is methodologically possible to formulate simultaneous models within this approach. The failure of both authors to achieve statistically and conceptually acceptable simultaneous models is more likely to be a function of the data available than to be a major methodological problem. No matter how convincing the statistical evidence may be, the final test of recursive models versus simultaneous models is their comparative predictive accuracy and ease of operation. Neither paper addresses these questions. Hence, I conclude that the matter of the preferred structure of travel demand models remains a matter for future research, preferably the near future.

REFERENCES

25. Kendall, M. G., and Stuart, A. The Advanced Theory of Statistics, 3rd Ed. C. Griffin and Co. Ltd., London, Vol. 1, 1969, p. 232.
26. Lianos, T. P., and Rauser, G. C. Distribution of Parameters in a Distributed Lag Model. Journal American Stat. Association, Vol. 67, No. 337, 1972, pp. 64-67.
27. Stopher, P. R., Spear, B. D., and Sucher, P. O. Toward the Development of Measures of Convenience for Travel Modes. Transportation Research Record 527, 1975.

David T. Hartgen, Planning and Research Bureau,
New York State Department of Transportation

These papers represent important research in disaggregate demand modeling and its application to transportation planning. The papers emphasize both the theoretical structure of these tools and model calibration using conventional travel data.

The previous discussions have concentrated on the differences in the findings of these studies, particularly the authors' results on the sequentiality or simultaneity of model structures. However, I am impressed more with the similarities between the authors' approaches and their general conclusions than with the differences in their empirical findings.

1. Both authors investigate in detail whether sequential or simultaneous structures are more valid for modeling travel behavior and rely heavily on multinomial disaggregate techniques for studying these problems.
2. Both authors conclude, generally speaking, that choice of a model has considerable influence over the coefficients that will be obtained and very likely the conclusions that may be drawn based on such models. Therefore, both authors call for great care in the selection of models and in their use.
3. With respect to data bases, both authors suggest the use of very small data sets, concentrating on the detail within records rather than the collection of large data bases. It is interesting to note that both research efforts were conducted with less then 150 observations, extremely small by current standards.
4. Both authors mention the aggregation problem as one that needs to be addressed as these models are applied in transportation planning.
5. Finally, both authors suggest a variety of detailed applications as a key element of further research.

The differences in the specific findings of these studies only serve to emphasize the authors' own caveats concerning the models and their use. We are dealing here with two very different kinds of problems, as the authors have pointed out. In Ben-Akiva's case a commonly studied travel component (off-peak shopping trips) is investigated within the context of several prior travel decisions (purpose and time of day). In Liou and Talvitie's case, a rather specialized problem in transportation planning (rail commuter trips) is further broken down for study into only the access portion. Given such differences, it is not surprising that the empirical findings of these studies are different. In fact, I would have been surprised had they been identical, given the great differences in the contexts being studied.

Of particular interest are the implications in this research for use of these tools in transportation planning. It seems that a number of events must occur before the tools described here will be included in the general repertoire of transportation planning procedures. The first of these is that the profession must know a great deal more about the kinds of transportation problems to which such tools can be logically applied. This is particularly true with complicated choice combinations, including purpose, time of day, route, mode, and destination choices. There are not many cases in which a planner would want to model that entire choice sequence as one set of simultaneous choices. Ben-Akiva's results, it seems, obtain partially because he chooses a problem that by its nature lends itself easily to the assumptions of the Luce model. On the other hand, Liou and Talvitie's findings probably stem from the fact that they study a problem drawn from the context of a broader but integral choice decision. It is doubtful whether either of these authors would have obtained the same results had they extended their choice contexts to, say, trip purpose choices in Ben-Akiva's case or primary mode choices in Liou and Talvitie's case. The models constructed by these authors have only been demonstrated and tested in problems involving logically paired choices and have not been extended to more complicated sequences.

This is a situation the profession will have to live with for some time. It appears unreasonable to expect that the simultaneous models suggested by Ben-Akiva will immediately be applicable to a broad range of intricate transportation choices. There are certainly choice contexts to which these models can be applied, but perhaps first these tools should be embedded within the overall multichoice transportation planning procedure currently in common use, perhaps replacing one or more of those steps.

This phase is probably required for practical reasons as well. Except for a few individuals primarily in the academic and consulting environments, transportation planners are not, generally speaking, well acquainted with the underlying theory and application of disaggregate techniques. By and large, the profession consists of individuals trained in conventional UTPS procedures. There is a general dissatisfaction with the conventional UTPS approach, but the disaggregate techniques are not perceived to hold the answer to these problems, at least not yet. What is required now are demonstrations of the application and potential savings of these procedures, with particular attention to computer processing and institutional constraints to model implementation. It is simply not an easy job to bring on line a disaggregate model and to make a case in an agency for its use, as opposed to a currently used conventional procedure. Parallel use of both tools is more feasible in the short run. In the interim, it seems probable that these procedures will remain relatively unused until the profession is more convinced of their utility.

Of particular concern here is the relevance of models to the profession in general. Most transportation planners are much more concerned with the usefulness of models to their work than they are with the theoretical niceties of their structure. If models do not have the right policy variables or cannot be used to address questions at issue today, then whether they are constructed by using disaggregate techniques or conventional aggregate procedures will be equally irrelevant, for in neither case will the models be used. Therefore I suggest that the most important step we can take to ensure that the results of research such as this will be used is to ensure that our models are capable of addressing relevant policy questions. This means not just the study of demand model structures, although that is important, but also the specification of appropriate variables and identification of particular problem contexts in which those variables can be used to predict behavior and to estimate impact. How many of us today can say we have on-line demand models capable of addressing questions related to the energy crisis, car pooling, pricing schemes, fuel policies, and parking modal interface? I doubt that many of us can.

AUTHORS' CLOSURE

Moshe Ben-Akiva

The discussions by Kraft, Stopher, and Hartgen deal with three important aspects of the proposed travel demand models: model structure and estimation procedure, empirical evidence, and applicability to transportation planning. Before discussing these topics, I will briefly restate my line of argument and in particular the precise purpose of the empirical study.

RESEARCH STRATEGY

The research begins with the assumption (which was not questioned by any of the discussants) that the choices of frequency, destination, mode, and time of day for a specific trip purpose (e.g., shopping) are elements in a single decision. In other words, it is assumed that a potential traveler compares alternative trips and therefore jointly selects a frequency, destination, mode, and time of day for a given travel purpose. Any sequence assumed for choices that are elements of a single decision is arbitrary. For some decisions, a conditional decision-making process implying a specific sequence of choices may be a realistic assumption. However, a joint decision-making assumption is more realistic for nonwork trips such as shopping trips.

Alternative assumptions about the decision-making process or about the causal relationships among choices result in different model specifications. A simultaneous structure is used to represent jointly determined choices, whereas a recursive structure is used to represent a specific choice sequence. In addition to differences in their mathematical formulations, the alternative models are estimated differently. In a simultaneous structure a model for the joint probability is estimated directly, whereas models for the conditional and marginal probabilities in a recursive structure are estimated separately.

Based on a priori reasoning, the simultaneous model is superior to a recursive model because it is a more realistic representation of the behavioral process. We do not expect these different models to produce identical estimation results. The value of empirical evidence is in determining the consequence or practical significance of the differences between the models. The differences are in the estimation results and potentially also in the ease and cost of estimation and application.

The empirical study that was conducted and reported in the paper was therefore not designed to test which model is better. Rather, it was designed to determine the feasibility of the simultaneous model and the practical significance of the differences in the estimation results between the alternative models. The empirical study indicated that a simultaneous model is feasible and that the differences in parameter estimates are of significant practical importance. Based on these results, it was recommended that travel demand models be developed by using the simultaneous model structure.

Since completion of this study, several other empirical studies have strengthened these conclusions (28, 30, 34). The results of Liou and Talvitie also support the conclusions that the estimation of a simultaneous model is feasible and that the alternative models produce differences in estimation results that are of practical importance. (The fact that the estimation results of their simultaneous model were not satisfactory can be attributed to a poor specification.)

MODEL STRUCTURE AND ESTIMATION PROCEDURE

Kraft states that models formulated with a recursive (or conditional) structure do not imply a sequential (or conditional) decision-making process but "are simply a factoring of a joint probability distribution into the product of a conditional and a marginal distribution." Furthermore, the fact that alternative model structures produced different estimation results leads Kraft to conclude that the results obtained are inconsistent with probability theory.

There are no inconsistencies between the results and probability theory. The differences between the simultaneous and the recursive models are the direct result of different mathematical formulations and different estimation procedures.

Any given model, whether simultaneous or recursive, can be expressed mathematically as a joint probability or as a sequence of marginal and conditional probabilities. However, the mathematical expression for a joint probability derived from a recursive model will, in general, be different from the formulation of the joint probability in a simultaneous model. Likewise, marginal probability derived from a joint model will in general be different from its specification in a recursive model. The reason for these differences is the need to introduce additional assumptions in a recursive model in order to formulate composite variables. This is the basic difference between simultaneous and recursive structures. An additional behavioral assumption of a sequence of choice is embedded within a composition rule. Thus, the two structures are mathematically different, and we should not expect identical results.

This is illustrated by using the logit model and the example of mode and destination choice as in my paper. We can estimate a single logit model that explains directly the joint probability of shopping destination and mode choice as follows:

$$P(d, m:DM) = \frac{e^{U_{dm}}}{\sum\limits_{d'm' \epsilon DM} e^{U_{d'm'}}} \tag{13}$$

This model treats the choices of mode and destination jointly (i.e., simultaneously), does not require any sequence assumptions, and allows for a realistic representation of choice between complete alternatives (e.g., shopping trip to the CBD by bus versus a car trip to a suburban shopping center).

Using the logit model for each choice separately in one specific sequence, we estimate the following two models:

$$P(d:D) = \frac{e^{U_d}}{\sum\limits_{d' \epsilon D} e^{U_{d'}}} \tag{14}$$

$$P(m:M_d) = \frac{e^{U_{m|d}}}{\sum\limits_{m' \in M_d} e^{U_{m'|d}}} \tag{15}$$

The basic difficulty with a recursive structure is the representation of variables that vary across more than one choice in the utility function of the marginal probability. Travel time, for example, varies between mode and destination combinations. Therefore, this variable enters directly into the joint utility U_{dm} in Eq. 13 and the conditional utility $U_{m|d}$ in Eq. 15, but it cannot be directly represented in the marginal utility U_d in Eq. 14 because the mode is indeterminate. Denoting this variable as X_{dm}, for a joint model we can write the following [for simplicity other variables in the utility functions are not explicitly included (2, ch. VI)]:

$$U_{dm} = U(X_{dm}) \tag{16}$$

where the travel time variable, for example, enters directly the utility function. For the sequential model, we can write the conditional utility as

$$U_{m|d} = U^m(X_{dm}) \tag{17}$$

and the marginal utility as

$$U_d = U^d(X_d) \tag{18}$$

where X_d is assumed to represent the values of X_{dm} by all alternative modes. Thus,

$$X_d = g(\{X_{dm}, \forall m \in M_d\}) \tag{19}$$

where g( ) is some composition rule and X_d is a composite variable of X_{dm} across modes. An example of such a definition is

$$X_d = \sum_{m \in M_d} X_{dm} \times P(m:M_d) \tag{20}$$

This is the rule that was used in the shopping model estimated by CRA (4). It implies a sequence assumption, and it requires that a lower level conditional probability (to compute these composite variables) be estimated before a higher level marginal probability can be estimated or predicted.

The assumption implied for the variable of travel time, for example, is that destination choice is based on expected travel time across modes. This means that the actual choice of mode is indeterminate when the destination choice is made; it is assumed that destination is chosen first, and then conditional on the destination alternative chosen a mode is chosen.

Thus, modeling a set of choices, which are realistically assumed to be made jointly, in a recursive structure with composite variables will result in errors due to model misspecifications.

The difference in estimation procedures between the simultaneous and the recursive structures also contributes to differences in the results. The differences can be attributed to both efficiency issues and specification errors. Consider a simultaneous model such as the one for the joint probability of mode and destination in Eq. 13. We can mathematically derive the expression for any conditional or marginal probability. Therefore, although the model is specified as simultaneous we can estimate the model coefficients in two ways. We can either estimate the joint probability directly or estimate any sequence of marginal and conditional probabilities, say P(d) and P(m|d). What are the differences between these two estimation procedures? [The answer to this question is discussed in detail by Ben-Akiva (2, ch. IV).]

First, the estimation results obtained by the direct estimation of the joint probability are more efficient than those obtained for the marginal and conditional probabilities. This is because all the data are used to estimate the coefficients appearing in $P(m|d)$ in the first case, whereas only the data on alternative modes to the chosen destination are used in the second case to estimate the same coefficients. In addition, the coefficients estimated for the conditional probability are used to create the composite variables used in the estimation of the marginal probability. Thus, there is also a propagation of errors in a sequential estimation, where the randomness in lower stage estimates shows up as measurement errors in the higher stage models. Thus, in the absence of specification errors, the difference in the estimation results can be explained by random variation. The direct estimation of the joint probability provides the most reliable estimates.

Second, the differences between the two estimation procedures are also attributed to specification errors. Specification errors will affect both estimation procedures. Because the effect of the error on the joint probability estimation will be different from that on a conditional and a marginal probability estimation, it contributes to the differences that will be observed between the two procedures. A more careful selection of variables may reduce the specification errors and therefore reduce the differences between the estimation results of the joint and conditional probabilities. However, even if the model is fully specified, the differences between simultaneous and recursive models that result from use of composite variables in a recursive structure will still be present.

To overcome these difficulties, Kraft suggests that constraints be included in the estimation process that will ensure consistency between different sequences. This suggestion is implemented in the direct estimation of the joint probability of a simultaneous model.

The logical answer to Kraft's question, Can we develop a model structure that does not do violence to our theories or to reality yet is both mathematically and statistically tractable? is the simultaneous model recommended in my paper. If we consider a wider range of travel-related choices, such as employment location, automobile ownership, and residential location in addition to short-run choices, the appropriate model structure may be termed block recursive. In this structure, the blocks of long-run and short-run choices are recursive with respect to each other. However, within each block the choices are made jointly and modeled in a simultaneous structure.

EMPIRICAL EVIDENCE

The essence of Stopher's discussion is that the conclusions are not supported by the supplied statistical evidence. However, as noted earlier, my conclusions are not based on hypothesis testing or solely on comparisons of goodness-of-fit measures. From a theoretical point of view the simultaneous model is superior to the recursive model. The empirical evidence is used to demonstrate the feasibility of the simultaneous model and to determine the practical significance of the differences that we expect a priori.

Stopher states that I did not establish that the differences of the estimates of the recursive and the simultaneous models are statistically significant. (It is not clear to me what specific statistical test can be used for this purpose because the models have different nonlinear mathematical formulations. The goodness-of-fit measures of the different models are almost identical.) Suppose for the moment that the differences are not statistically significant at a reasonable significance level. Should this be a reason to revise my conclusions? One knows a priori that there are differences due to different mathematical formulations and estimation procedures; therefore, it can be concluded that the lack of significant statistical differences is due to the small sample size. Furthermore, the coefficient estimates of the simultaneous model are more reliable because of the direct estimation of the joint probability. Inasmuch as a simultaneous model does not cost much more to estimate and apply than a recursive model, which has been indicated, the simultaneous model structure should be recommended.

Stopher further questions the empirical evidence due to the use of so-called engineering rather than perceived choice sets and attribute values. He goes so far as to say that using engineering data in model estimation "provides no behavioral information and consigns the exercises to academic esoterica." Alternatively, the use of perceived values can be considered the academic esoterica, inasmuch as the models are intended to provide decision-makers with forecasts of the impacts of alternative policies or plans. A prerequisite for the use of a model estimated with perceived (or reported) values for forecasting is a set of relationships between perceived and engineering values. Furthermore, perceptions are only intermediate variables formed by individuals on the basis of physical objects and characteristics that are measured by the engineering values. Thus, a model that uses engineering estimates explains directly the individual's reaction to physical objects and characteristics, circumventing the need to deal with the intermediate variables of perceptions. It is clear, however, that the model functional form and parameters embody both the formation of perceptions and the behavioral response to these perceptions. This rationale is the basis for a large body of econometric literature, and I am not aware of any studies that have shown it to be "behaviorally incorrect" as Stopher claims.

I agree with Stopher's statement that a lack of constants in a choice model specification presupposes a fully specified model. Constants can be excluded from the model only if it can be shown that they have no explanatory power; i.e., the constants are equal to zero. (The specification used by Liou and Talvitie did not include constants. This may explain their unsatisfactory estimation results for the simultaneous model.) This was not done in this study. The constants included are a mode-specific constant, DA, and a CBD destination constant, DCBD. It does not make sense to use a constant for every possible destination because there are so many.

In summary, the conclusion that the simultaneous structure is preferred does not depend on additional statistical evidence. Future effort should be directed at the steps necessary to make fully specified simultaneous travel demand models available for application rather than at further comparisons of alternative model structures.

APPLICABILITY TO TRANSPORTATION PLANNING

The applicability of the proposed models to transportation planning, the essential issue of research in travel demand modeling, is the focus of Hartgen's discussion. I agree with Hartgen's conclusion that "what is required now are demonstrations of the application and potential savings of these procedures, with particular attention to computer processing and institutional constraints to model implementation." Furthermore, the overall direction for the development of travel demand models should be toward operational models that are more reliable in a forecasting context.

Several studies that have been completed since my paper was written have further shown the practicality of estimating simultaneous choice models in situations with more than two dimensions of choice and with a very large number of alternatives and observations. First, the simultaneous model of destination and mode for the shopping trip has been extended to include frequency as well in the simultaneous structure (29). A similar simultaneous model was also applied successfully with a data set from the Netherlands (30). Simultaneous disaggregate choice models have also been successfully applied to the automobile ownership and work mode choices (31) and are currently being applied to the entire set of long-run locational and automobile ownership choices (33).

From the point of view of aggregate forecasting the use of simultaneous models places no additional burden on data requirements and computational efforts. Two ongoing research efforts at M.I.T. focus on the application of simultaneous travel demand models to aggregate forecasting. The first study is investigating alternative procedures of using these models for aggregate predictions (32). The second study (sponsored by DOT University Grant Research Program) is implementing these models and procedures for a transportation planning case study.

REFERENCES

28. Adler, T. SM thesis in progress, Department of Civil Engineering, M.I.T., Cambridge.

29. Adler, T., and Ben-Akiva, M. A Joint Frequency, Destination and Mode Choice Model for Shopping Trips. M.I.T., Cambridge, 1974.
30. Buro Goudappel en Coffeng B. V. and Cambridge Systematics, Inc. Disaggregate and Simultaneous Travel Demand Models: A Dutch Case Study. Deventer, The Netherlands, and Cambridge, Massachusetts, 1974.
31. A Behavioral Model of Automobile Ownership and Mode of Travel. Cambridge Systematics Inc., 1974.
32. Koppelman, F. PhD dissertation in progress, Department of Civil Engineering, M.I.T., Cambridge.
33. Lerman, S. PhD dissertation in progress, Department of Civil Engineering, M.I.T., Cambridge.
34. Nerlove, M., and Press, S. Univariate and Multivariate, Log, Linear and Logistic Models. Rand Corporation, Santa Monica, California, 1973.

Peter S. Liou and Antti P. Talvitie

The issues raised by the discussants are well taken, and perhaps we can clarify some of the questions that have been brought up.

It is true that mathematically the joint probability should have the same value as the product of a conditional probability and a marginal probability and be independent of any specific mathematical formulation; e.g., $P(s, m) = P(s|m) \times P(m) = P(m|s) \times P(s)$. However, the conditional probability, by definition, means the probability of one event taking place given that the other event has already taken place. Therefore, from the behavioral viewpoint, decisions of the access mode and station choices may be approached from three directions: the simultaneous decision-making process and the two sequential decision-making processes. It appears in this study that empirically it is possible to estimate choice models based on the station-mode sequential assumption. Further research is necessary to determine whether the joint probabilities would indeed be the same had it been possible to estimate choice models based on the other sequential structure or based on the simultaneous modeling structure.

With regard to the modeling aspect of this study, the importance of including one or more constants in the utility function was investigated. In fact, in the various access mode and station models, mode-specific dummy variables were included to indicate the access mode with which the in-vehicle time was associated. Nevertheless, all the models estimated in this fashion involved (significant) coefficients with incorrect signs. One such model is in Table 16 (model 3).

Another point concerning the modeling aspect is the access mode alternatives. Although it was observed in the survey whether a traveler drove or was driven to a station, the detail of the data did not permit separating these two modes. Consequently, the automobile mode was viewed as a single mode and the value of each level-of-service variable associated with the automobile mode was obtained by averaging the values by automobile-driver mode and automobile-passenger mode (22).

In formulation of the station selection models, differences in train fare among alternative stations were not included because the alternative stations were generally adjacent to each other and were located within the same fare zones.

With regard to the model evaluation aspect of this study, two points need to be mentioned. As reported in the paper, no valid models could be obtained by using the simultaneous model structure. The estimated models were considered invalid on the basis of incorrect coefficient signs. The coefficients with wrong signs were statistically significant; therefore, further testing between the models was not carried out. For the simultaneous model, for example, the standard error of the bus time variable is 0.0427, whereas the variable coefficient is 0.1974. The second point concerns the standard error of the implied value of time. The variances of the implied values of time were determined in the same way as suggested in one of the discussions (19). It was not known, however, whether their distributions were normal or skewed.

Another important aspect discussed concerns the applicability of the multinomial logit modeling technique to transportation planning and forecasting. This is a laudable goal for transportation research. However, it seems that the objective is not only to replace the existing UTPS package with a new package but to replace it with a better and more valid modeling system or concept. Therefore, it ought to be realized that transportation planning should be done on the basis of methodology and assumptions that are consistent with travel choice behavior and consumer theory, especially when such planning involves both short-range predictions and long-range projections extending 20 years into the future. This would require a full and extended knowledge and understanding of the extent and limitations of this and other modeling techniques.

In closing, we would like to emphasize that formulation of the access mode and station choice models in this study was based on the assumption that the modeling of the access trip can be separated from the rest of the journey. Further research is necessary to determine whether this consideration is proper. On the other hand, the behavioral decision-making process involved in making a trip is so complex that structuring a single simultaneous model to include all the trip choices such as purpose and household location on the one hand and trip frequency, time of day, trip destination, travel mode, and route on the other is clearly not advisable, if not impossible. Some assumptions have to be made to separate the various trip choices and thus simplify the modeling process. This study only explores one of the many such assumptions that could be made.

Finally, it is stressed in all the discussions that further research in this area is warranted. As pointed out, knowledge of the relevant alternatives as perceived by individual travelers and specific socioeconomic information (e.g., individual income, household expendable income) are important in formulating behavioral models. Unfortunately, these types of information were not available for this study. Furthermore, the validity and the effects of a number of assumptions that are essential for disaggregate multinomial modeling such as the separability assumption and of course the axiom of independence of irrelevant alternatives should be investigated in the near future.

THE NATURE OF TRAVEL DECISION-MAKING

Reid H. Ewing, Department of City and Regional Planning, Harvard University

Travel decision-making is described in behavioral terms, and an alternative to conventional travel forecasting is suggested. Two issues are considered. The first is the order of travel decision-making. The second is the interaction of travel decisions. This paper defines the order of travel decision-making with expected utility theory, and it describes the interaction of travel decisions with dynamic programming. The resulting travel model is based on theories of decision-making and is unique in this respect.

•TRANSPORTATION planning has been dominated by engineering and economics throughout its brief history. As a result, conventional travel forecasting neglects the human element in intraurban trip-making. It is possible to improve travel forecasting by first defining the nature of travel decision-making and then applying the result to travel demand modeling. That is the purpose of this study.

Two issues are addressed: In what order are travel decisions made, and in what way do travel decisions interact? The literature on decision-making under uncertainty provides an answer to the first question. And the literature on dynamic decision-making provides an answer to the second. These answers lay the foundation for a unique travel demand model based on current theories of human behavior. This model is proposed as a practical alternative to existing travel demand models.

THE APPEAL OF SEQUENTIAL TRAVEL MODELS

In a recent paper, Brand (4) described alternative methods of travel demand modeling. His alternatives included sequential and simultaneous models. This paper is limited to sequential travel models for two reasons. First, they are more efficient than simultaneous models because the number of travel options (each of which must be evaluated) increases multiplicatively when decisions are combined. Two departure times for each of two modes, each with 10 alternative routes to 10 destinations, translate into 400 ($2 \times 2 \times 10 \times 10$) travel options for each origin in a simultaneous travel model. With literally hundreds of origins and additional times, modes, routes, and destinations, simultaneous travel demand models become unwieldy. Second, travel decisions are apt to be made sequentially rather than simultaneously. The multitude of options available to travelers forces them to simplify the decision-making process. They do this by making sequential decisions, thereby greatly reducing the number of travel options they must consider. Just as travel demand is modeled sequentially in response to the limitations of the digital computer, so may travelers simplify the decision-making process in response to their own limitations. Arguments of this nature are persuasive enough to justify the present emphasis on sequential travel models. Simultaneous travel models and the empirical choice between simultaneous and sequential models will be the subject of future research.

EXPECTED UTILITY THEORY APPLIED TO TRAVEL DECISION-MAKING

Conventional travel forecasting assumes that travelers choose a time of departure, a destination, a mode, and a route in that order. In some travel models, the choice of mode precedes the choice of destination, but the order is always predefined and invariant. This can be criticized on two counts. First, the conventional order is based on neither theory nor observation. It would certainly be fortunate if the conventional order proved to be correct. Second, it is likely that the order of travel decision-making depends on the available alternatives and thus varies with the situation. Any model that assumes a fixed order of travel decision-making cannot consistently reproduce the decision-making process. The following discussion is motivated by these concerns.

The order of travel decision-making has been ignored in the literature with the exception of a paper by Brand (4). Brand suggests several alternative methods of ordering travel decisions: order based on (a) information, decision-making proceeds from the most informed to the least informed decisions; (b) adjustment, decision-making proceeds from the least easily adjusted to the most easily adjusted decisions; and (c) timing, decision-making proceeds from the latest to the earliest decisions in time (that is, the logical order of decisions runs counter to their sequence in time).

It is not possible to choose among the three methods with the limited evidence available [see Feger and Feger (12) and Tate and Howell (23)]. Fortunately, no choice is necessary because a single theory incorporates them all. This theory is well established empirically and ranks among the foremost theories of decision-making under uncertainty.

Few subjects in psychology have received more attention than decision-making under uncertainty. Two reviews of this subject are noteworthy. Edwards (9) provides a useful introduction to the subject. Luce and Suppes (16) go into much greater detail. Readers who desire additional information should consult these reviews.

A prominent theory of decision-making under uncertainty, expected utility theory, defines the order of travel decision-making. Expected utility theory states that decision-makers, when faced with uncertainty, make decisions that maximize their expected utility. That is, they select the option with the greatest expected utility, where the expected utility of an option is given by

$$EU = \sum_{n} u_i p_i \tag{1}$$

where

u_i = utility of outcome i,
p_i = probability that outcome i will occur, and
n = number of possible outcomes.

Consider a decision-maker with several options. His first option has two possible outcomes. One outcome has a utility of 10 and occurs 40 percent of the time, and the other has a utility of 5 and occurs 60 percent of the time. From Eq. 1, the expected utility of this option is 7 ($10 \times 0.4 + 5 \times 0.6$). The decision-maker will choose this option only if his other options offer less expected utility.

Travel decision-making necessarily involves uncertainty if it is a sequential process. This is true even if individual travel decisions are made under certainty because future decisions are unknown. For example, if their first travel decision is the choice of departure time, travelers' neglect of destinations, modes, and routes introduces uncertainty into decision-making. Only after all travel decisions are made can the choice of departure time be evaluated with certainty.

Travel decision-making is comparable to gambling. Each decision represents a gamble. It is assumed that travelers evaluate all four travel decisions and make the decision with the greatest expected utility; then they reevaluate the remaining travel decisions and make the decision with the greatest expected utility. This continues until they have made all travel decisions.

Travelers are assumed to have complete knowledge of alternative departure times, destinations, modes, and routes. Individual decisions are therefore made under certainty, and the following equation applies (15, 22):

$$p_i = u_i \Big/ \sum_{n} u_i \tag{2}$$

where p_i is the probability that alternative i will be chosen, u_i is the utility of alternative i, and the summation is taken over all n alternatives.

Combining Eqs. 1 and 2 gives an expected utility of

$$EU = \sum_n u_i^2 \Big/ \sum_n u_i \tag{3}$$

Equation 3 defines the order of travel decision-making. Let us assume that a departure time and destination have been chosen already. Either a mode is chosen next and then a route or vice versa. It will be shown that the order of travel decision-making depends on (a) the number of alternative modes and routes and (b) the similarity of alternative modes and routes. Consider the following examples.

In the first example, there are more alternative routes than modes, but alternative modes and routes are equally similar. The utility of mode and route pairs is

$$\begin{aligned} M_1R_1 &= 11 = R_1M_1 \\ M_1R_2 &= 8 = R_2M_1 \\ M_1R_3 &= 5 = R_3M_1 \\ M_2T_1 &= 7 = R_1M_2 \\ M_2R_2 &= 4 = R_2M_2 \\ M_2R_3 &= 1 = R_3M_2 \end{aligned}$$

M_1 refers to mode 1, M_2 to mode 2, R_1 to route 1, and so on. The expected utility of travel is 7.50 if the choice of mode precedes the choice of route and 7.46 if the choice of route precedes the choice of mode. These values are obtained from Eq. 3 by first calculating the expected utility prior to the last decision and then the expected utility prior to the next to last decision. This indicates that travel decisions with few alternatives are made before travel decisions with many alternatives.

In the next example, alternative routes are more similar than alternative modes, but the number of alternative modes and routes is the same. The utility of mode and route pairs is

$$\begin{aligned} M_1R_1 &= 10 = R_1M_1 \\ M_1R_2 &= 8 = R_2M_1 \\ M_2R_1 &= 4 = R_1M_2 \\ M_2R_2 &= 2 = R_2M_2 \end{aligned}$$

The expected utility of travel is 7.56 if the choice of mode precedes the choice of route and 7.62 if the choice of route precedes the choice of mode. This indicates that travel decisions with similar alternatives are made before travel decisions with dissimilar alternatives.

The conclusions of the last two paragraphs may surprise readers. Intuitively, decisions with many dissimilar alternatives should precede decisions with few similar alternatives. Yet these examples indicate that the reverse is true. There is nothing inconsistent about this. Just as the expected utility of travel is calculated by evaluating the last decision first and just as Brand's third method of ordering travel decisions assumes that the logical order of decisions runs counter to their sequence in time, travel decision-making may begin with the last travel decision and work backward.

It is encouraging that order based on information and order based on adjustment represent special cases of the present theory. Incomplete information about travel alternatives results in uncertainty in decision-making and causes the probabilities of selection to converge. In the limit, if no information is available, the probability of selection is the same for every alternative. This decreases the expected utility of corresponding travel decisions and causes travelers to postpone these decisions until other decisions are made. Order based on information applies in this case. When travel decisions are long-lived, the probabilities of selection are affected, but in this case they tend to diverge rather than converge. It is likely that travelers exercise greater care when they make long-lived decisions, and therefore they select alternatives with maximum utility. This increases the expected utility of long-lived travel decisions and causes travelers to advance these decisions relative to other travel decisions. Order based on adjustment applies in this case.

DYNAMIC PROGRAMMING APPLIED TO TRAVEL DECISION-MAKING

Conventional travel demand models assume that travel decisions are made independently, whereas direct demand travel models assume that travel decisions are fully integrated. It is likely that travel decisions are neither so independent nor so fully integrated as assumed in these models.

In his review of sequential decision-making, Edwards (8) divided sequential decisions into six classes. His sixth class included dynamic decisions of the type made by travelers, which are characterized by the dependence of later decisions on earlier decisions. Because travel decision-making is a dynamic process, the literature on dynamic decision-making applies and a substantial body of knowledge is available. During the 1950s Bellman (2) developed an analytical technique known as dynamic programming. At that time Bellman suggested that dynamic programming could be used to simulate dynamic decision-making. Dynamic programming has since been applied to dynamic decision-making in a variety of theoretical and empirical studies (5, 14).

Several studies have compared the performance of subjects on dynamic decision-making tasks to the results of dynamic programming. Using variations of the Reader's Control Problem to simulate recurrent business decisions, Rapoport and Ray applied dynamic programming to stochastic problems (19), adaptive problems (18), adaptive problems of unknown duration (20), and deterministic problems (21). In all cases dynamic programming adequately described subjects' performance on dynamic decision-making tasks.

Dynamic programming is based on the principle of optimality, which states that the best decision at each stage in the decision-making process is the decision that optimizes the remainder of the process. Dynamic programming therefore begins with the desired objective (a maximum benefit from travel) and works backward through the sequence of decisions to the starting point (the decision to work, shop). Each decision in the sequence is optimized according to a predefined decision rule. A decision-maker can ignore past and future decisions and evaluate his present alternatives with this decision rule.

In many ways, travel decision-making is an ideal problem for dynamic programming. Problems must be divided into stages in dynamic programming. Travel decision-making has four stages—the choice of departure time, destination, mode, and route. Alternative states must be defined at each stage. Alternative departure times, destinations, modes, and routes represent alternative states of the various travel decisions. An objective function must be optimized in some manner. The objective function in travel decision-making is the utility of travel and, of course, it is maximized. Actions and policies must be defined. Choices among alternative departure times, destinations, modes, and routes represent actions, and sets of choices represent policies. Returns on all actions and policies must be evaluated. This presents a problem because no utility accrues in travel until all travel decisions are made. However, it may be possible to associate utility with individual travel decisions. The utility of individual travel decisions will be a function of independent measures of performance (i.e., measures that do not depend on other travel decisions) such as distance between zones in trip distribution, smoothness of ride in modal split, and directness of route in network assignment.

Once the utility of individual decisions is determined, a dynamic programming model can be developed. The following notation is used:

$f_n(a)$ = maximum utility of all remaining travel decisions if alternative a is chosen in the nth decision (= objective function in state a and stage n if decision-making is optimal);

u_{ab} = utility of travel alternative b in the n+1 stage of travel decision-making if travel alternative a was chosen in the nth stage (= the return from the action of choosing state b in stage n+1 when state a was chosen in stage n);

t = alternative departure times;

d = alternative destinations;

m = alternative modes;

r = alternative routes; and

o = state of travelers after travel decision-making.

If travel decisions are made in the conventional order (i.e., departure time, destination, mode, and route), the following equations can be derived from the principle of optimality:

$$f_4(r) = \max_{u_{ro}} [u_{ro}]$$

$$f_3(m) = \max_{u_{mr}} [u_{mr} + f_4(r)]$$

$$f_2(d) = \max_{u_{dm}} [u_{dm} + f_3(m)]$$

$$f_1(t) = \max_{u_{td}} [u_{td} + f_2(d)]$$

By way of example, the third equation says that the maximum utility of all remaining travel decisions if destination d is chosen in the second stage of travel decision-making is equal to the maximum value of the sum of the utility of destination d plus the maximum utility of all remaining travel decisions if mode m is chosen in the third stage of travel decision-making. $f_2(d)$ is evaluated for all modes that serve destination d. Modes that do not are ignored. The other equations are analogous.

Dynamic programming identifies the optimal departure time, destination, mode, and route. It does not distribute trips among alternative departure times, destinations, modes, and routes as conventional models do. Distribution is a separate process. Several methods of distribution are available. The most promising is an intervening opportunities approach that assigns a constant fraction of all remaining trips to the best of the remaining travel alternatives (the process is then repeated without this alternative).

The use of dynamic programming to simulate travel decision-making has one major drawback. Dynamic programming simulates optimal decision-making, but travel decision-making is apt to be suboptimal. The issue in travel forecasting is not how travelers should make decisions but how travelers do make decisions. Arriving at the same conclusion (for decision-making in general), Rapoport extended his earlier analysis to include suboptimal decision-making (17). He began by noting that shortcomings of the human memory, the cost of information-gathering and -processing, the inability to plan ahead, the ignorance of interdependencies, and so on limit our ability to make decisions. Decision-making "in non-trivial tasks will, in general, not be optimal." He went on to propose a theory of dynamic decision-making that makes use of programming algorithms. Decision-makers are assumed to plan ahead one, two, or even more decisions in dynamic decision-making tasks. The extent to which they plan ahead depends on their ability and the nature of their tasks. Rapoport described empirical tests of his theory. Subjects' performance on Elithorn's perceptual maze test (10) was compared to the results of three algorithms. The first algorithm assumed that subjects plan ahead one move when choosing among alternative paths through a maze, the second that they plan ahead two moves, and the third that they plan ahead three moves. The performance of many subjects corresponded to one of the three algorithms (particularly to the third one) and varied with the design of the maze. It would appear that planning horizons do vary from individual to individual and from task to task.

If travel decision-making is suboptimal, Rapoport's algorithms can be used in travel forecasting. One algorithm applies to travelers with a planning horizon of one travel decision, another to travelers with a planning horizon of two travel decisions, and a third to travelers with a planning horizon of three travel decisions. (A planning horizon of four travel decisions leads to optimal decision-making.) Algorithms can be chosen by comparing the results of each to the behavior of travelers. The algorithm that best describes the travel behavior of each socioeconomic class can be used in forecasting.

It should be noted that existing travel models correspond to the extremes of travel decision-making. Conventional travel demand models assume that travel decisions are independent of each other. Trip generation, trip distribution, modal split, and network assignment are modeled independently and undertaken sequentially. Each is completed without reference to the others. Conventional travel forecasting corresponds to the simplest type of travel decision-making, where travelers ignore all future decisions. The result of sequential decision-making is identical to that of dynamic decision-making if the planning horizon of travelers is one travel decision. In contrast, direct demand models assume that travel decisions are fully integrated. Trip generation, trip distribution, modal split, and network assignment are combined in a single model. Because all stages are undertaken simultaneously, they are allowed to fully interact and influence each other. Direct demand forecasting corresponds to the most sophisticated type of travel decision-making, where travelers consider all travel decisions simultaneously. The result of simultaneous decision-making is identical to that of dynamic decision-making if the planning horizon of travelers includes all remaining travel decisions. Existing models can describe the extremes of travel decision-making, but only behavioral models can describe travel decision-making in general.

Transportation planning has been dominated by engineering and economics throughout its history. Hopefully, this paper and another by the author (11) have demonstrated the potential of behavioral models in transportation planning. I believe that transportation planning is ready to incorporate psychological theory into its simple economic models.

ACKNOWLEDGMENTS

I gratefully acknowledge the assistance of Daniel Brand of Harvard University who suggested I restructure this study along the present lines and critiqued the first draft of this manuscript.

REFERENCES

1. Abelson, R. The Choice of Choice Theories. In Decision and Choice (Messick, S. and Brayfield, A. H., eds.), McGraw-Hill, 1964.
2. Bellman, R. Dynamic Programming. Princeton University Press, 1957.
3. Brand, D. Theory and Method in Land Use and Travel Forecasting. Highway Research Record 422, 1973.
4. Brand, D. Travel Demand Forecasting: Some Foundations and a Review. HRB Spec. Rept. 143, 1973.
5. Ebert, R. J. Sequential Decision Making: An Aggregate Scheduling Methodology. Psychometrika, Vol. 36, 1971, pp. 303-316.
6. Edwards, W. A Bibliography of Research on Behavioral Decision Processes to 1968. Report 7, University of Michigan, Ann Arbor, 1969.
7. Edwards, W. Bibliography: Decision Making. Engineering Psychology Group. University of Michigan, Ann Arbor, 1964.
8. Edwards, W. Dynamic Decision Theory and Probabilistic Information Processing. Human Factors, Vol. 4, 1962, pp. 59-73.
9. Edwards, W. Behavioral Decision Theory. Annual Review of Psychology, Vol. 12, 1961, pp. 473-498.
10. Elithorn, A., Jagoe, J. R., and Lee, D. N. Simulation of a Perceptual Problem-Solving Skill. Nature, Vol. 211, 1966, pp. 1029-1031.
11. Ewing, R. H. Psychological Theory Applied to Mode Choice Prediction. Transportation, to be published.
12. Feger, H., and Feger, B. Contributions to the Content Analytical Investigation of Decisions. Archiv für die gesamte Psychologie, Vol. 121, 1969, pp. 205-254.
13. Fishburn, P. C. Decision and Value Theory. Wiley and Sons, 1964.
14. Kaufmann, A. The Science of Decision-Making. McGraw-Hill, 1968.
15. Luce, R. D. Individual Choice Behavior. Wiley and Sons, 1959.
16. Luce, R. D., and Suppes, P. Preference, Utility and Subjective Probability. In Handbook of Mathematical Psychology, (Luce, R. D., Bush, R. R., and Galanter, E., eds.), Vol. 3, Wiley and Sons, 1965.
17. Rapoport, A. Optimal and Suboptimal Decisions in Perceptual Problem-Solving Tasks. Behavioral Science, Vol. 14, 1969, pp. 453-466.

18. Rapoport, A. Dynamic Programming Models for Multistage Decision-Making Tasks. Journal of Mathematical Psychology, Vol. 4, 1967, pp. 48-71.
19. Rapoport, A. A Study of Human Control in a Stochastic Multistage Decision Task. Behavioral Science, Vol. 11, 1966, pp. 18-32.
20. Rapoport, A. A Study of a Multistage Decision-Making Task With an Unknown Duration. Human Factors, Vol. 8, 1966, pp. 54-61.
21. Ray, H. W. The Application of Dynamic Programming to the Study of Multistage Decision Processes in the Individual. Ohio State University, Phd. thesis, 1963.
22. Restle, F. Psychology of Judgment and Choice: A Theoretical Essay. Wiley and Sons, 1961.
23. Tate, J. D., and Howell, W. C. Term Expectation and Uncertainty in Human Decision Behavior. U.S. Air Force Aerospace Medical Research Laboratory, TDR 63-118, 1964.

STRUCTURAL MODEL FOR EVALUATING URBAN TRAVEL RELATIONSHIPS

Edward J. Kannel, University of Illinois at Urbana-Champaign; and
Kenneth W. Heathington, University of Tennessee

Urban travel forecasting equations are typically developed by analyzing only the relationships between several possible explanatory variables and the ultimate variable of interest, trip production. Seldom is the full degree of interaction among explanatory variables such as automobile ownership, household size and income, and accessibility understood. In this paper, a structural model is used to examine the relationships among an entire system of variables rather than just the simple isolated effects. The basic concepts and limitations of the modeling approach are discussed, and models of urban household trip production are evaluated. Several conclusions about the causal structure of urban travel relationships are drawn. The structural model is felt to be an important methodological tool for developing urban transportation theory.

•THE AVAILABILITY of multiple linear regression computer programs has made the linear regression model a popular technique for estimating travel demand in urban transportation planning studies. The user can handle large quantities of data and develop models without fully understanding the procedure or assumptions of the regression model. Evaluation of the model is based more on statistical goodness of fit than on an understanding of the causal structure that exists among the variables examined.

In a traditional trip generation model, evaluation of the relationships among a set of variables would be desirable. A typical trip generation model may use automobile ownership and household size to forecast home-based trip production. However, before auto ownership can be used, a forecast of this variable must be developed. Average family size and income might be used as the explanatory variable to estimate car ownership. These two models do not allow examination of the entire structure of the relationships that exist among family size and income, auto ownership, and trip production.

A complementary analysis technique is discussed that can be used to evaluate the direct effect of income on trip production as well as the indirect effect of income on car ownership rates. The approach discussed has alternatively been referred to as causal models, structural models, path models, and recursive models. These terms will be used interchangeably.

A structural model is a system of equations that allow the analyst to evaluate fully the interrelationships among a system of variables. To be sure, multiple regression analysis allows the analyst to observe the effects of several independent variables, either alone or in combinations, but this is possible only for the single response variable under immediate consideration. Although cross-product terms of selected independent variables may be included to show the effects of interaction among the independent variables, neither the nature of this interaction nor the relative contribution of the component parts can be evaluated. The structural model, on the other hand, uses a set of equations that outline the causal priorities of the variables and permit predictions of how a change in any variable in the system affects other intermediate variables in the system as well as the dependent variable of interest.

DEVELOPMENT AND APPLICATION OF STRUCTURAL MODEL

Pioneer work in the analysis of path coefficients, the parameters of a structural model, was done by Wright in the area of genetics (13). Blalock, drawing on the writings of Simon (10) and Wold (12), provided the major thrust for a study of causal inferences in nonexperimental research of the social sciences (1). Duncan (3), Land (8), and Heise (4) described their studies of causal models and outlined systematic approaches for interpretation of the model.

Analysis of structural models has also received attention in the development of transportation forecast models. Kain used a system of recursive equations to evaluate the interrelationships among variables that affect work trip length (5). He hypothesized a four-stage recursive model in which the decision process is such that the worker first selects an environment in which he wishes to live (space preference). This choice is influenced by factors such as sex, income, and housing prices. The first decision will then affect the choice of car ownership, which in turn affects the travel mode choice. Finally, all three of these affect the length of the work trip. In any of these equations additional variables may contribute to the variation in the dependent variables; however, these variables were taken to be exogenous to the system, and no attempt was made to define the interrelationships among exogenous variables.

More recently, de Neufville and Stafford summarized the difficulties in interpretation that arise when strictly correlational models are used rather than when the causal structure among a set of variables is examined (2). They also evaluated travel demand through structural model studies.

BASIC CONCEPTS OF CAUSAL MODELS

The goal of causal modeling is to develop a set of relationships that correspond to real-world causal processes. Analysis of a causal structure requires specification of a network of causal paths that exist between the variables of interest and identification of the parameters of causation so that the effects of each variable on the other variables in the model can be measured. The mathematical equations that make up the causal structure are a set of recursive equations. [For more detailed discussion of the computational procedures, readers are referred to Heise (4) and Land (8). Also, Blalock (1) offers extensive arguments on the need for caution in making causal inferences and clarifies the conditions under which causal inferences may be possible.] The following four-variable system is an example of a set of recursive equations (Fig. 1):

$$Z_1 = P_{1a}Z_a$$

$$Z_2 = P_{2b}Z_b$$

$$Z_3 = P_{31}Z_1 + P_{3c}Z_c$$

$$Z_4 = P_{41}Z_1 + P_{42}Z_2 + P_{43}Z_3 + P_{4d}Z_d$$

The structural model is composed of three types of variables. Exogenous variables are considered input to the system. It is assumed that these variables are completely determined by other variables outside of the system and that neither the nature of their origin nor the correlation that may exist between these inputs is of concern for the model being considered. Paths between the exogenous variables are represented in the figure as two-headed curvilinear arrows. These paths indicate only that some correlation exists. No direction of causality is assumed. In Figure 1, Z_1 and Z_2 are the exogenous variables of the system.

Measured variables within the structural model are endogenous variables. Unlike exogenous variables, the total variation in the endogenous variables is of interest. The total variation in the endogenous variables is assumed to be completely determined by some linear combination of exogenous variables, other endogenous variables, and some unmeasured residual or error variable. The postulated causal relations among the variables are shown in Figure 1 by unidirectional arrows extending from each determining variable to each variable dependent on it. In the models discussed, it is assumed that there is only one direction of causation; i.e., if X causes Y, Y cannot in turn be a cause of X. Variables Z_3 and Z_4 are endogenous.

Because it is unrealistic to assume that the variation of a system variable can be determined completely by other measured variables of the system, residual variables are introduced. It is assumed that residual variables are uncorrelated with the set of variables immediately determining the variable under consideration and that they have a mean value of zero.

Residual variables are represented on a path diagram by unidirectional arrows. The residual variable paths have alphabetic subscripts to distinguish them from the paths of the measured variables, which have numerical subscripts. For simplification, the residual paths and the paths between the exogenous variables are often eliminated from the causal model diagram. Variables Z_a, Z_b, Z_c, and Z_d are the residual variables in Figure 1.

The exogenous variables, Z_1 and Z_2, are assumed to be completely determined by outside forces Z_a and Z_b, which are either unknown or just not of interest in the analysis. The path coefficients P_{1a} and P_{2b} are equal to one and are not normally included in the model or diagram.

Estimation of the model parameters for each equation revolves around fitting a model to the data so that the amount of variation contributed by the residual variables is minimal. The parameters of the structural model are computed by using the least squares criteria common to linear regression analysis; however, the relative importance of the determining variables is interpreted by using standardized regression coefficients. This standardized parameter or path coefficient P_{ij} is a measure of the fraction of the standard deviation of the dependent variable for which the independent variable is directly responsible. More definitely, it is the fraction that is found if the factor varies to the same extent as in the observed data, all other variables being constant.

The relationship between the regular regression coefficient b_{ij} and the path coefficient P_{ij} is

$$P_{ij} = b_{ij} \frac{\sigma_j}{\sigma_i}$$

where σ_i and σ_j are the standard deviations of the dependent and independent variables. The path coefficients are also referred to as beta coefficients.

Use of the standardized parameter facilitates the computations necessary to evaluate how consistently the model reproduces the interrelationships that exist in empirical data. Using standardized coefficients, we can show that, for the model given in Figure 1, the structural model predicts the following linear correlation between variables Z_4 and Z_1:

$$r'_{41} = P_{41} + P_{42}r_{12} + P_{43}P_{31}$$

where

r'_{41} = predicted correlation for the hypothesized model;
r_{12} = observed correlation between the exogenous variables; and
P_{41}, P_{42}, and P_{31} = path coefficients estimated in regression analysis (8).

The total correlation between Z_4 and Z_1 is composed of three elements. First, there is a direct effect between the variables indicated by path coefficient P_{42}. Secondly, there is an indirect effect, $P_{43}P_{31}$, caused by the influence that variable Z_4 has on Z_3, which in turn influences Z_4. Finally, there is another indirect effect, $P_{42}r_{12}$, that encompasses the correlational effect of the exogenous variables. It must be cautioned that these direct and indirect effects can be interpreted only for the model under study. The direct effect is a true, isolated direct effect only if the other independent variables are orthogonal to the variable being considered and if the effects of all other variables are removed. In practice, these conditions are seldom completely met.

The entire correlation matrix could be reproduced in like manner by using the additional relations derived for the model in Figure 1:

$$r'_{13} = P_{31}$$

Figure 1. Example of four-variable causal model.

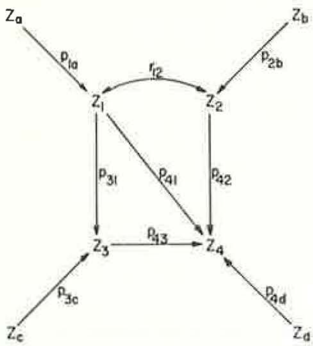


Figure 2. Causal ordering of household travel relationships based on the first model.

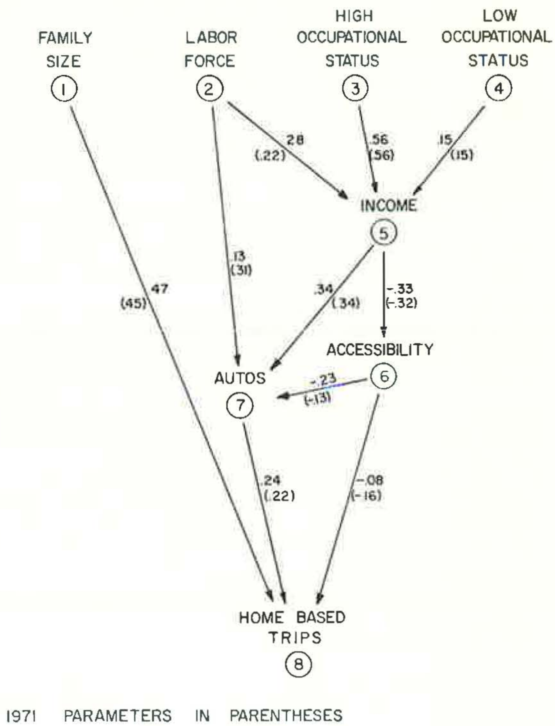


Table 1. Observed correlation matrix of causal model variables.

Variable	Family Size	Labor Force	High Occupation	Low Occupation	Income	Accessibility	Automobiles	Home-Based Trips
1964								
Family size	1.00	0.24	0.01	0.10	0.08	-0.07	0.16	0.52
Labor force		1.00	0.07	0.14	0.34	-0.03	0.26	0.31
High occupation			1.00	-0.83	0.46	-0.23	0.25	0.06
Low occupation				1.00	-0.28	0.15	-0.18	0.05
Income					1.00	-0.33	0.47	0.29
Accessibility						1.00	-0.35	-0.20
Automobiles							1.00	0.35
Home-based trips								1.00
1971								
Family size		0.41	0.12	0.20	0.15	-0.11	0.37	0.55
Labor force			0.16	0.32	0.36	-0.10	0.44	0.33
High occupation				-0.65	0.50	-0.27	0.25	0.24
Low occupation					-0.15	0.10	0.08	0.06
Income						-0.32	0.49	0.31
Accessibility							-0.27	-0.27
Automobiles								0.42
Home-based trips								

$$r'_{23} = P_{31}r_{12}$$

$$r'_{24} = P_{42} + P_{41}r_{12} + P_{43}P_{31}r_{12}$$

$$r'_{34} = P_{43} + P_{41}P_{31} + P_{42}P_{31}r_{12}$$

If a variable explains a significant portion of the variance in a dependent variable, it would exhibit a strong direct effect in the causal model and would logically be considered an important element of the model. However, if a particular variable does not show a strong direct contribution in a given equation, the analyst would not immediately reject the variable from the entire structure. Instead, the importance of the variable in the other equations of the system would be considered. If the variable is significant in other relationships it would be an important variable in the overall system. The advantage of the structural model is that all relationships can be examined simultaneously and the faithfulness of that system in reproducing the empirical relationships in the data set can be evaluated.

EVALUATING THE MODEL

The purpose of developing a causal model is to help the analyst understand the relationships among a set of variables that are important in some behavioral process. An objective in testing these relationships is to obtain a model that adequately reproduces the conditions that occur in empirical data and yet is as parsimonious as possible. For any postulated model, one can compute the correlation matrix and compare this with an observed correlation matrix. In the least parsimonious case in which all possible paths are included, there are no conditions imposed on the model to test the adequacy of that model. Any ordering of the variables results in a reproduced correlation matrix that exactly equals the empirical correlations (8). In this case only the knowledge of the causal priorities determines selection of one model over the other. The problem then is to make an initial determination of the significance of the paths in the model.

If sufficient data are available, analysis of variance models may serve as a starting point for evaluating the significance of variables that might be introduced in the model. The analysis of variance provides a measure of significance not only of the main effects of the variables but also of possible interactions. Because interaction terms are not included in the simplified model, there may be incorrect interpretations from the structural model.

A second method for eliminating paths in the causal models is to retain only those variables that are statistically significant according to the F-test criterion used in regression analysis. However, as the sample size becomes large, path coefficients that make very small contributions to the total variance may be judged statistically significant and retained in the model. For this situation, Land suggests that the analyst choose a minimum value below which a path coefficient is considered substantively insignificant.

Finally, when an over-identified model has been structured, (i.e., a model in which one or more possible paths have been eliminated) additional constraints will be established. A model is judged adequate if it reproduces correlations between the system variables in accordance with the imposed constraints. If these predicted correlations adequately represent the empirical correlations, the analyst might accept this as the best representation of the causal structure or he might check the possibility of eliminating other paths. If the model is not adequate, the analyst either reverts to the previously accepted model or tests some other model in which a different link is eliminated.

APPLICATION TO URBAN TRANSPORTATION PLANNING

The data analyzed in this paper were collected as part of a research project developed to study the temporal stability of household trip production (6, 7). The data were obtained from home interview surveys in Indianapolis. The first survey was conducted in 1964 as part of a traditional urban transportation planning study. The second survey was conducted in 1971. The same families were interviewed in both studies so that variation in household travel behavior, which may be due to family preferences, type of dwelling unit, or location within the urban structure, would be minimized. This sample of households was specifically selected to represent all levels of three principal socioeconomic variables, i.e., family size, automobile availability, and income; however, this simultaneously provided a wide range of other characteristics such as occupational status, educational levels, and location of residence within the urban area.

Although travel relationships at the household level of analysis do not have the apparent statistical strength of those obtained from zonal averages, it was hypothesized that relationships existing at the household level have greater behavioral validity and causal significance and therefore are more temporally stable than zonal model analyses. The stability of these household relationships over a 7-year period was examined in this study.

One of the causal model relationships of interest in this study was an evaluation of the hypothesis that the trip production from households is affected by the accessibility of the household to major activity centers within the urban area. The measures of accessibility used in the study were those developed by Nakkash for the 1964 highway network (9). Relative accessibility variables were determined from the friction factors of the calibrated gravity model. The relative accessibility variable was therefore a function of trip purpose and auto travel time, but was weighted by the amount of a given activity in a zone. Nakkash developed relative accessibility measures for several activity types such as employment, retail floor area, and school floor area, but these are all intercorrelated and only employment accessibility is used in the discussion. This variable tends to decrease with increasing travel time from the central city. Also it must be noted that accessibility measures comparable to the 1964 study could not be reproduced in 1971 because a complete new transportation study was not being conducted. The relative accessibility of each household in 1971 was therefore assumed to be equivalent to the 1964 calibrated values. All statistical tests were performed on the 1964 data.

STRUCTURE OF THE CAUSAL MODEL

The ordering of the causal network was based on a priori knowledge of the variables under consideration and on previous research models (5). Although several simplified models were tested to evaluate items such as the direct and indirect influence of income on trip production, only two models are discussed. The hypothesized formulation is a four-stage recursive model. The model hypothesizes that a family chooses a residential location based on a desire for a certain life-style, quality and style of housing, and preference or need for more or less space. Differences in preferred housing conditions may be shaped by factors such as individual attitudes, age, stage in the family life cycle, and family size. A family that needs more space would tend to locate in lower density areas that are relatively less accessible to major activity centers. The ability to satisfy the desire for housing type and space consumption, however, is controlled by the ability of the family to pay for the desired living style. Thus, income level of the family must be considered. Income might be determined by several factors such as education, occupation, age, and number of working members in the household.

Once the decision about housing requirements and residential location is made, the level of available transportation from that location influences the level of car ownership. Families living in high-density neighborhoods with greater accessibility may have the opportunity to satisfy some of their transportation needs by use of public transportation. Further, because of greater accessibility, more travel needs such as school or social-recreational trips are satisfied by the walking mode. The level of auto ownership may also be affected by family size, labor force, household income, and social status of the family.

Finally, the trip production rate of the household may be affected by any of the variables mentioned. The task is to evaluate and understand the degree of influence a change in one variable has on other variables in the model.

The causal ordering assumed in this study is shown in Figure 2. Only a small number of possible exogenous variables were actually considered in the model. Further, many of the possible linkages that could be included in the model have been eliminated because they were found to be of little importance as explanatory variables.

The correlation matrix for the variables is given in Table 1. The exogenous variables were family size, labor force, and occupation of the head of household. Occupation of the head of household was stratified into three groups and used as a dummy variable in the analysis. The groupings were nongainful, high status, and low status. The high status group was composed of professionals, managers, and salesmen; the low status group contained all other employed individuals. The nongainful dummy class was omitted from the analysis to allow solution of the least squares equations (11).

ANALYSIS OF THE MODEL

Examination of Figure 2 shows that many of the path coefficients are very stable for the 7-year period. The ability of the model to reproduce the correlations that exist among all of the variables can be evaluated by examining Table 2. If the model adequately represents all the existing relationships, these differences should approach zero. Although several of the possible links have been removed, the model does reproduce the correlation matrix quite well. The major discrepancy occurs in its ability to reproduce the relationship between family size and auto ownership in the 1971 data set. The calculated correlation was 0.20, whereas the observed correlation was 0.37. The changing relationship between family size and auto ownership may be due to the maturation process that has occurred. As the families moved through stages of the life cycle and children from the larger families moved out of the household, the relationship between family size and car ownership stabilized. As a result, the linear correlation between these variables increased substantially from 0.16 to 0.37. The same basic change was noted between labor force and auto ownership.

The model was examined for other possible links that might be removed to make the model more parsimonious. Earlier analysis suggested that the effect of accessibility on home-based travel is indirect because of its association with auto ownership (6). A two-way analysis of variance of the 1964 data set indicated that, when ownership and accessibility were tested, ownership was the significant variable and no interaction was found. The path coefficient in the model (-0.08) also indicates that the direct path is small and explains only a small portion of the variance in travel.

A second model, in which the direct link from accessibility to home-based travel has been removed, is shown in Figure 3. The differences in observed and empirical correlations are given in Table 3. The ability of the structural equations of the second model to reproduce the empirical correlation matrix is essentially the same as that of the first model. Only the relationship between trip production and accessibility is altered by removing this causal link. Because new measures of accessibility were not available in 1971, we could not determine whether the difference in the correlation in 1971 was a function of nonmeasurement error or actual changes in the effect of accessibility over time. Because the analysis of variance of the 1964 data found accessibility to be insignificant, the second model was chosen as the final structural model. This model was accepted as the most plausible explanation of the causal relationships among the variables.

SUMMARY

The causal relationships investigated in this study were restricted to models that meet the following assumptions.

1. A change in the dependent variable always occurs as a linear function of changes in the determining variables, and the effects of all other variables are assumed to be held constant.
2. The system contains no reciprocal causations, i.e., the model is strictly a recursive system. Reciprocal causations cause problems in identification. Although methods of treating such models are available, the procedure is more complex and the interpretations are clouded.
3. The causal priorities are sufficiently well known so that the structure of the

Table 2. Differences between empirical and reproduced correlations based on the first model.

Variable	Family Size	Labor Force	High Occupation	Low Occupation	Income	Accessibility	Automobiles	Home-Based Trips
1964								
Family size	—[a]	—[a]	—[a]	—[a]	0.00	-0.04	0.09	0.02
Labor force		—[a]	—[a]	—[a]	0.00	0.08	-0.02	0.12
High occupation			—[a]	—[a]	0.00	-0.08	0.05	-0.01
Low occupation				—[a]	0.00	0.05	-0.08	0.03
Income					—[a]	0.00	0.00	0.10
Accessibility						—[a]	0.01	-0.02
Automobiles							—[a]	0.04
Home-based trips								—[a]
1971								
Family size					-0.04	-0.05	0.17	0.04
Labor force					0.00	0.01	0.00	0.03
High occupation					0.00	-0.10	0.01	0.11
Low occupation					0.00	0.06	0.03	-0.03
Income						0.00	0.00	0.07
Accessibility							0.04	-0.02
Automobiles								0.08
Home-based trips								

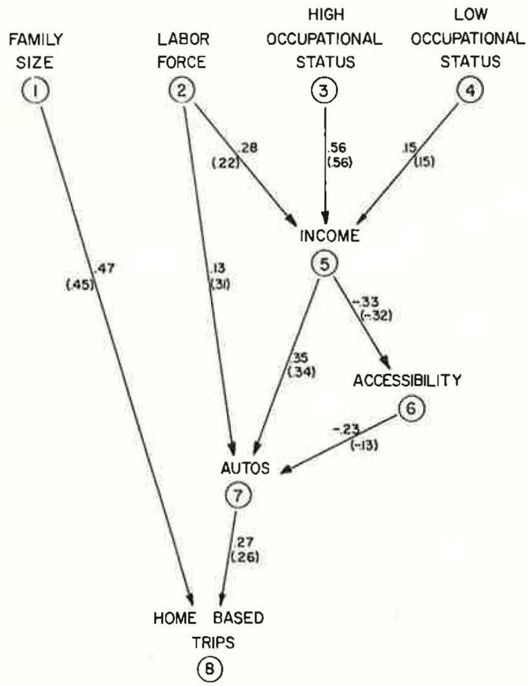
[a]These differences, by definition, must be zero.

Table 3. Differences between observed and reproduced correlations based on the second model.

Variable	Family Size	Labor Force	High Occupation	Low Occupation	Income	Accessibility	Automobiles	Home-Based Trips
1964								
Family size	—[a]	—[a]	—[a]	—[a]	0.00	-0.04	0.09	0.02
Labor force		—[a]	—[a]	—[a]	0.00	0.08	-0.02	0.12
High occupation			—[a]	—[a]	0.00	-0.08	0.05	0.00
Low occupation				—[a]	0.00	0.05	0.08	0.03
Income					—[a]	0.00	0.00	0.12
Accessibility						—[a]	0.01	-0.09
Automobiles							—[a]	0.04
Home-based trips								—[a]
1971								
Family size					-0.04	-0.05	0.17	0.04
Labor force					0.00	0.01	0.00	0.03
High occupation					0.00	-0.11	0.00	0.12
Low occupation					0.00	0.06	0.03	-0.04
Income						0.00	0.00	0.10
Accessibility							0.00	-0.17
Automobiles								0.08
Home-based trips								

[a]For the model specified, these differences must be zero.

Figure 3. Causal ordering of household travel relationships based on the second model.



1971 PARAMETERS IN PARENTHESES

model can be established as the correct ordering of the variables. It is not necessary that all correct paths are known, but the order of causation should be clear.

4. The data are generally measured on an interval or ratio scale, but, as in regular regression analysis, dummy variables may be used if caution is exercised in the interpretation of the results.

5. The usual assumptions of multivariate regression analyses are met.

The analysis model is a simplification of the real-world system and allows the analyst to evaluate the relative direct and indirect effects of the variables within the system. The model examined here is simplistic from the practical standpoint, for not all determinants of travel could be included. The model is simplistic in the statistical sense in that only linear relationships are considered and no interaction terms are specifically introduced for consideration.

Variables that have been found significant in household trip generation analyses were evaluated. The inferences obtained from the analysis indicate that auto ownership and family size have the most direct influence on trip generation rates. Income and level of accessibility to activity centers in the urban area also have an impact on travel, but this influence is indirect because of their influence on auto ownership. As household income increases the families tend to live farther from central employment concentrations and thus have lower relative accessibility to employment. In turn, these households exhibit higher auto ownership rates. The number of household members in the labor force also is a determinant of auto ownership rates, both directly and through its corresponding relationship with household income.

Finally, the effect of occupational status of the head of the household can best be understood by the extent to which it affects household income. Occupational status is an important variable for explaining variations in income, but the direct effects on any other variables in the causal model are negligible.

Although the analysis of causal inferences is subject to practical limitations, when properly used and interpreted it can provide a methodological tool for theory development.

ACKNOWLEDGMENTS

The authors are grateful to the Joint Highway Research Project of Purdue University, the Indiana State Highway Commission, and the Federal Highway Administration for the financial support that made this research possible. The opinions, findings, and conclusions expressed are those of the authors and not necessarily those of any of the supporting agencies.

REFERENCES

1. Blalock, H. M., Jr. Causal Inferences in Nonexperimental Research. University of North Carolina Press, 1964.
2. de Neufville, R., and Stafford, J. H. The Identification of Transportation Demand. In Urban Transportation Innovation (Brand, D., ed.), American Society of Civil Engineers, 1971.
3. Duncan, O. D. Path Analysis: Sociological Examples. American Journal of Sociology, Vol. 72, 1966.
4. Heise, D. R. Problems in Path Analysis and Causal Inferences. In Sociological Methodology (Borgatta, E. F., and Bohrnstedt, G. W., eds.), Jossey-Bass Inc., San Francisco, 1969.
5. Kain, F. J. A Multiple Equation Model of Household Locational and Trip Making Behavior. Rand Corporation, Santa Monica, Memo. 3086, 1962.
6. Kannel, E. J. A Disaggregate Analysis of Urban Travel Behavior. Purdue University, Joint Highway Research Project Rept. 17, July 1972.
7. Kannel, E. J., and Heathington, K. W. Temporal Stability of Trip Generation Relations. Highway Research Record 472, 1973.
8. Land, K. Problems in Path Analysis and Causal Inferences. In Sociological Methodology (Borgatta, E. F., and Bohrnstedt, G. W., eds.), Jossey-Bass Inc., San Francisco, 1969.
9. Nakkash, T. Z., and Grecco, W. L. Activity-Accessibility Models of Trip Generation. Highway Research Record 392, 1972.
10. Simon, H. Models of Man. John Wiley and Sons, 1957.
11. Suits, D. The Use of Dummy-Variables in Regression Equations. American Statistical Association Journal, Vol. 52, Dec. 1957.
12. Wold, H. J. L. Demand Analysis. John Wiley and Sons, 1953.
13. Wright, S. The Method of Path Coefficient. Annals of Mathematical Statistics, Vol. 5, Sept. 1934.

A COMPARATIVE EVALUATION OF INTERCITY MODAL-SPLIT MODELS

John C. Bennett, Raymond H. Ellis, and John C. Prokopy,
Peat, Marwick, Mitchell and Company; and
Melvyn D. Cheslow, U.S. Department of Transportation

Proposals for improved intercity transportation service, ranging from improved high-speed rail service to short take-off and landing or restricted take-off and landing air service, have been advanced for many intercity corridors in the United States. Transportation planners have been called on to forecast patronage and revenue of new transportation services and the diversion from existing services. Frequently, however, new travel surveys or model development is not possible, and reliance must be placed on existing models and secondary data sources. This paper provides a comparative evaluation of seven intercity modal-split models that have been developed in the last 5 years and recommends models for application in intercity transportation sketch planning analyses. The models discussed are evaluated in terms of their ability to replicate the observed travel patterns and in terms of their implied elasticities for the rail mode. Model CN27 was selected as the best overall model. It is stratified by purpose, which creates a data requirement that cannot always be reliably fulfilled. Thus, unstratified model CN22, second best among those tested, is recommended for use when base year travel data on trip purpose are unavailable.

•THE SEVEN intercity modal-split models considered in this paper are all calibrations of the cross-elasticity model (2, 3, 4, 5). This model has the following formulation:

$$S_i = \frac{C_i \prod_j (X_{ij})^{\alpha_{ij}}}{\sum_i \left[C_i \prod_j (X_{ij})^{\alpha_{ij}} \right]} \tag{1}$$

where

i = index identifying a mode,
j = index identifying a modal attribute,
X = transportation network variable,
S = variable identifying modal split, and
C_i, α_{ij} = calibrated coefficients.

Equation 1 is calibrated on a base mode (generally automobile). The ratios of the share of each mode i to that of the base mode (mode m) are considered as follows:

$$\frac{S_i}{S_m} = \frac{C_i \prod_j (X_{ij})^{\alpha_{ij}}}{C_m \prod_j (X_{mj})^{\alpha_{mj}}} \tag{2}$$

In turn, the logarithmic form of Eq. 2 is

$$\mu_i = r_i + \sum_j \alpha_{ij} v_{ij} - \sum_j \alpha_{mj} v_{mj}$$

Table 1. Calibrations of cross-elasticity model.

Model	Description	Remarks
CN22	Abstract mode, 1969	Calibrated early in the NEC Project; rail coefficient was lowered to improve model during NEC work
CN25	Abstract mode, 1969 (6)	Basis for analysis presented in first NECP report (7)
CN26	Abstract mode	
CN27	Stratified, abstract mode, 1971	Used for some analyses in second NECP report (8)
CN28B	Stratified, abstract mode, 1971	Used for some analyses in second NECP report (8)
HSGT	Abstract mode, 1972 (9)	Used in High Speed Ground Transportation Alternatives Study in 1972 (9)
SRI	Abstract mode, 1971 (10, 11)	

Table 2. Parameters for cross-elasticity model calibrations.

Model	Mode	C	a_1	a_2	a_3	k
CN22	Air	1.01	-2.23	-1.11	0.53	0.12
	Rail	1.46	-2.23	-1.11	1.05	0.12
	Bus	0.83	-2.23	-1.11	0.05	0.12
	Automobile	1.0	-2.32	-1.16	0	0
CN25	Air	1.1144	-1.9102	-0.9551	0.3247	0.12
	Rail	1.1144	-1.9102	-0.9551	0.3247	0.12
	Bus	1.1144	-1.9102	-0.9551	0.3247	0.12
	Automobile	1.000	-1.9288	-0.9644	0	0
CN26	Air	1.8978	-1.9135	-0.8555	0.5536	0.007
	Rail	3.8547	-1.9135	-0.8555	0.5536	0.007
	Bus	1.4486	-1.9135	-0.8555	0.5536	0.007
	Automobile	1.0	-1.9135	-0.8555	0	0
CN27						
Business	Air	1.1232	-3.384	-0.483	2.279	0.12
	Rail	1.4813	-3.384	-0.483	2.279	0.12
	Bus	0.3767	-3.384	-0.483	0	0
	Automobile	1.0	-3.384	-0.483	0	0
Nonbusiness	Air	0.7767	-1.5821	-1.5821	2.0462	0.18
	Rail	1.9881	-1.5821	-1.5821	2.0462	0.18
	Bus	1.3872	-1.5821	-1.5821	0	0
	Automobile	1.0	-1.5821	-1.5821	0	0
CN28B						
Business	Air	0.937	-3.384	-0.483	5.587	0.50
	Rail	1.2368	-3.384	-0.483	5.587	0.50
	Bus	0.3767	-3.384	-0.483	0	0
	Automobile	1.0	-3.384	-0.483	0	0
Nonbusiness	Air	1.1163	-1.5821	-1.5821	5.587	0.672
	Rail	1.4710	-1.5821	-1.5821	5.587	0.672
	Bus	0.9324	-1.5821	-1.5821	0	0
	Automobile	1.0	-1.5821	-1.5821	0	0
HSGT	Air	1.90	-1.9135	-0.8555	0.5536	0.007
	Rail	1.90	-1.9135	-0.8555	0.5536	0.007
	Bus	1.135	-1.9135	-0.8555	0	0
	Automobile	1.00	-1.9135	-0.8555	0	0
SRI	Air	1.50	-1.5	-1.5	0.3247	0.12
	Rail	0.75	-1.5	-1.5	0.3247	0.12
	Bus	0.75	-1.5	-1.5	0.3247	0.12
	Automobile	1.00	-1.5	-1.5	0	0

where

$$\mu_i = \log \frac{S_i}{S_m}$$

$$r_i = \log \frac{C_i}{C_m}$$

$$V_{ij} = \log \chi_{ij}$$

Calibrations of the cross-elasticity model considered in this paper are given in Table 1. The specification for these calibrations is as follows:

$$S_i = \frac{w_i}{\sum\limits_i w_i}$$

$$w_i = Ct^{a_1} c^{a_2} f'^{a_3}$$

$$f' = (1 - e^{-kf})$$

where

t = total average one-way door-to-door travel time in hours,
c = total average one-way door-to-door travel price in dollars, and
f = average number of daily one-way departures in one direction.

(The automobile per-person price is 1 cent per mile for CN22, CN25, and nonbusiness trips for CN27 and CN28B. The price is 1.2 cents per mile for CN26 and HSGT and is 2.3 cents per mile for business trips for CN27 and CN28B. The price for SRI is 1.76 cents per mile.)

The calibrated parameters for the models are given in Table 2.

STRUCTURING THE ANALYSIS

Data for city pairs in and outside the Northeast Corridor (NEC) were used to test the modal-split models. Each set of data consisted of annual volumes and measures of service attributes for each of the four modes serving the city pair.

A list of NEC city pairs that were considered best for testing the models was developed. Each potential city pair for which data were available was examined for the reliability of the modal travel volume information and the uniqueness of the impedance measures. City pairs involving Trenton were generally eliminated, for example, because of the lack of a clear-cut air service alternative. A traveler could use relatively poor service at Trenton Airport or drive for more than an hour and use good service at Newark Airport. A test data set consisting of data for 64 city pairs in the NEC was assembled.

Travel volume data for 44 city pairs outside the corridor were assembled from previous surveys (12, 13). Detailed modal travel volumes were available for 22 city pairs, whereas reliable secondary source travel volume data existed for the other 22.

The following criteria were used to evaluate the models:

1. Total number of trips by mode,
2. Root mean square error (RMSE) between the observed and the estimated modal trips,
3. Correlation between observed and estimated modal trips,
4. Slope of a linear regression fitted between the estimated modal trips (dependent variable) and the observed modal trips (independent variable), and
5. Intercept of the linear regression.

EVALUATION OF MODELS WITH NEC DATA

A comparison of the models with respect to the NEC data is given in Table 3. A comparison of the overall accuracy of the models suggests that, with some notable exceptions, all of the models exhibit similar error tendencies. Each of the models underestimates the total number of automobile trips. Bus travel is overestimated and air travel is underestimated by all models except CN25. Five of the models overestimate and three underestimate rail travel. The correlation between observed and estimated trips by automobile and air is above 0.9 for all models, and the correlation between observed and estimated rail travel is approximately 0.9 for all models, whereas the correlation for bus travel is generally lower than 0.9. This result and the fact that the ratio of RMSE to aggregate trips is high for rail and bus travel for all models indicate that the models generally estimate bus and rail travel more poorly than they estimate automobile and air travel.

Each model produces a positive value for the intercept of the linear regression [estimated = f (observed)] for all modes, which indicates that the models tend to overestimate modal travel for low-volume modal trip interchanges. The slopes of the linear regressions are less than one for automobile and air travel, the two largest segments of the intercity travel market, for all models. This suggests that the models tend to compensate for the overestimation at low volumes by underestimating automobile and air travel at higher volumes. In general, the variation in observed versus estimated modal volumes is sufficiently high to invalidate any conclusions drawn strictly on the basis of the regression parameters.

Table 4 gives a ranking of the models according to their ability to replicate modal totals (1 is best; 7 is worst). The models perform relatively unevenly among modes. CN28B ranks first for automobile, bus, and rail trip totals but last for air travel. CN27 is best for air travel but fourth for the other modes. CN22 and CN27 are the most consistent, for they are the only models that rank among the top four for all modes.

A comparison of the weighted average RMSE (weighted by observed modal split) and the RMSE for each mode estimated by each model is given in Table 5. The average ranking of all models is almost identical to the ranking for automobile travel because the automobile captures approximately 70 percent of the travel market in the test data set and the values of RMSE are higher for the automobile than for the other modes. CN27, a stratified model calibrated with the most recent data prior to the effort undertaken for this project, is ranked first overall. CN22, ranked third overall, is the most consistent for the four modes.

EVALUATION OF MODELS WITH NON-NEC DATA

The modal-split models were applied to the data for the 44 city pairs outside the Northeast Corridor, and the modal estimates were compared to observed travel volumes. The results are given in Table 6. In general, model performance was similar to that obtained by using the NEC data. Total automobile travel was underestimated by all models, and common carrier travel was overestimated in all but two cases.

The models may be ranked in order of prediction accuracy for modal trips (Table 7). The HSGT model is most accurate for rail, least accurate for bus, sixth for automobile, and second for air. This characteristic of the HSGT model is not unexpected, for it is an adjusted version of CN26 used for forecasting travel patronage for candidate high-speed ground transportation systems outside the NEC; the rail forecasting accuracy of the HSGT model is consistent with its principal application. Models CN22 and CN27 are relatively consistent.

The rank of each model on the basis of the RMSE measure for each mode is given in Table 8. Model CN22 is best overall. Model CN27 provides the least amount of variability between observed and estimated rail volumes and is ranked third for both automobile and bus; it ranks a relatively poor sixth, however, for air travel. As was the case for the test using NEC data, CN22 is the most consistent model among the modes, ranking in the top four for all modes.

Table 3. Summary statistics for model comparison using NEC data.

Mode	Statistic	Observed	CN22	CN25	CN26	CN27	CN28B	HGST	SRI
Automobile	Trips[a]	40.79	39.08	36.85	36.32	38.50	39.20	36.99	38.53
	RMSE[b]		178.0	215.6	213.9	175.7	176.7	218.0	194.4
	r		0.98	0.97	0.98	0.98	0.98	0.98	0.98
	Slope		0.92	0.86	0.83	0.89	0.91	0.85	0.90
	Intept[b]		27.2	28.1	39.4	36.6	34.2	36.6	31.5
Bus	Trips[a]	3.28	5.04	7.19	4.77	5.16	3.83	9.62	7.56
	RMSE[b]		70.4	125.6	83.8	72.7	51.8	185.1	136.9
	r		0.82	0.80	0.77	0.81	0.80	0.79	0.82
	Slope		1.21	1.74	1.27	1.22	0.93	2.22	1.92
	Intept[b]		17.0	22.9	9.4	18.1	12.3	36.7	19.8
Air	Trips[a]	6.18	5.77	6.48	5.60	5.99	7.08	5.53	5.95
	RMSE[b]		75.7	88.1	69.4	56.4	81.1	70.5	93.0
	r		0.96	0.91	0.97	0.97	0.91	0.97	0.92
	Slope		0.66	0.63	0.68	0.78	0.76	0.68	0.59
	Intept[b]		26.7	39.9	21.3	18.6	37.7	20.5	36.2
Rail	Trips[a]	8.15	8.52	7.89	11.71	8.74	8.29	6.27	6.36
	RMSE[b]		157.5	183.2	178.9	144.2	160.8	193.9	202.2
	r		0.91	0.90	0.88	0.92	0.91	0.89	0.92
	Slope		0.65	0.54	0.85	0.72	0.62	0.50	0.45
	Intept[b]		50.6	54.8	75.3	44.8	50.0	33.8	41.7

[a]Millions of annual trips. [b]Thousands of annual trips.

Table 4. Ranking of models according to their ability to replicate modal totals (NEC data).

Model	Automobile	Bus	Air	Rail
CN22	2	3	4	3
CN25	6	5	3	2
CN26	7	2	5	7
CN27	4	4	1	4
CN28B	1	1	7	1
HSGT	5	7	6	6
SRI	3	6	2	5

Table 5. Ranking by root mean square error between observed and estimated trips (NEC data).

Model	Automobile	Bus	Air	Rail	Average
CN22	3	2	4	2	3
CN25	7	5	6	5	6
CN26	5	4	2	4	5
CN27	1	3	1	1	1
CN28B	2	1	5	3	2
HSGT	6	7	3	6	7
SRI	4	6	7	7	4

Table 6. Summary statistics for model comparisons using non-NEC data.

Mode	Statistic	Observed	CN22	CN25	CN26	CN27	CN28B	HSGT	SRI	PML
Automobile	Trips[a]	35.76	33.59	30.96	33.97	33.83	33.51	31.06	32.83	26.30
	RMSE[b]		133.2	234.1	137.8	137.8	152.5	242.7	181.0	569.8
	r		0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98
	Slope		0.94	0.86	0.92	0.92	0.91	0.85	0.90	0.61
	Intept[b]		1.8	3.2	23.1	22.7	23.5	12.9	15.1	104.0
Bus	Trips[a]	1.31	2.72	3.85	1.69	2.69	1.95	5.78	3.26	10.57
	RMSE[b]		82.4	124.5	44.0	82.1	57.6	197.6	103.0	437.0
	r		0.63	0.70	0.74	0.62	0.63	0.69	0.69	0.72
	Slope		1.26	2.01	0.99	1.21	0.93	2.83	1.71	6.21
	Intept[b]		24.2	27.6	8.8	24.9	16.7	47.2	23.3	55.3
Air	Trips[a]	2.36	2.88	3.32	2.72	3.10	3.70	2.43	2.39	1.71
	RMSE[b]		56.7	71.6	62.3	87.1	97.5	64.0	65.1	122.8
	r		0.86	0.78	0.82	0.77	0.75	0.80	0.81	0.19
	Slope		0.73	0.70	0.71	0.94	0.98	0.61	0.54	0.14
	Intept[b]		26.4	37.8	23.8	19.8	31.7	22.3	25.6	31.1
Rail	Trips[a]	0.58	0.82	1.88	1.64	0.39	0.84	0.73	1.53	1.44
	RMSE[b]		24.7	69.0	60.1	16.9	38.0	19.7	52.9	76.4
	r		0.83	0.80	0.82	0.86	0.88	0.84	0.74	0.86
	Slope		1.13	2.15	2.03	0.68	1.71	0.94	1.57	2.68
	Intept[b]		3.8	14.4	10.5	-0.1	-3.4	4.3	14.0	-2.6

[a]Millions of annual trips. [b]Thousands of annual trips.

Table 7. Ranking of models according to their ability to replicate modal totals (non-NEC data).

Model	Automobile	Bus	Air	Rail
CN22	3	4	4	3
CN25	7	6	6	7
CN26	1	1	3	6
CN27	2	3	5	2
CN28B	4	2	7	4
HSGT	6	7	2	1
SRI	5	5	1	5

Table 8. Ranking by root mean square error between observed and estimated trips (non-NEC data).

Model	Automobile	Bus	Air	Rail	Average
CN22	1	4	1	3	1
CN25	6	6	2	6	6
CN26	2	1	2	3	2
CN27	3	3	6	1	3
CN28B	4	2	7	4	4
HSGT	7	7	3	2	7
SRI	5	5	4	5	5

Figure 1. Elasticity with respect to rail time.