HIGHWAY RESEARCH RECORD

Number 446

Travel Behavior

6 reports
prepared for the
52nd Annual Meeting

Subject Area

84 Urban Transportation Systems

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D.C. 1973

NOTICE

The studies reported herein were not undertaken under the aegis of the National Academy of Sciences or the National Research Council. The papers report research work of the authors that was done at the institutions named by the authors. The papers were offered to the Highway 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 Highway Research Board.

Before publication, each paper was reviewed by members of the HRB 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 Highway 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.

ISBN 0-309-02177-4
Library of Congress Catalog Card No. 73-11612
Price: $2.20

Highway 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 Highway Research Board, National Academy of Sciences, 2101 Constitution Avenue N.W., Washington, D.C. 20418.

CONTENTS

FOREWORD

The papers contained in this RECORD focus on microanalysis of urban transportation demand and modal-choice models that attempt to combine various socioeconomic variables with behavioral data.

Aldana, de Neufville, and Stafford describe a disaggregate or microanalytic demand model for urban travel. Two special features characterize the structure of the model. First, the household is taken as the basic decision unit. Second, it is proposed that the social class and the stage of the life cycle of these units are important explanatory variables.

Inglis discusses research into the use of a logical mathematical formulation to model modal split. The model investigates the extension to the multimodal situation in terms of both user and system variables. Time and cost difference is the major variable introduced for the systems; other variables are income, age, and a rush-hour dummy. The trip being modeled was a short commuter trip, which was only a part of a longer overall trip, and involved the access to a commuter rail station only. The line-haul portion was not considered.

Reichman states that travel time savings are usually estimated on the basis of objectively measured times. Such objective measures are assumed to correspond to the mean of a distribution of subjective time savings as reported by travelers. In a multimodal, single-route, interurban passenger survey carried out in Israel, this was not found to be the case. Discussion of the implication of these results and further suggestions for clarification of the findings are suggested by the author.

Watson reports the results of some empirical prediction tests on disaggregate, behavioral, stochastic models of modal choice. The author reports that the expected-number calculation (summation of probabilities) yields predictions of modal choice of a high degree of accuracy, although a caveat is in order regarding the transferability of such models: They should be transferred with care and only to similar situations.

Hall and Surti explain modal-choice factors and attitude patterns resulting from their research work in the Denver region.

Wigner presents the results of the calibration of modal-choice models designed to be used as a part of the urban transportation planning package for the Chicago area and also as regional planning and policy tools by themselves.

MICROANALYSIS OF URBAN TRANSPORTATION DEMAND

Eduardo Aldana, University of Andes, Bogota , Columbia;
Richard de Neufville, Massachusetts Institute of Technology; and
Joseph H. Stafford, University of Florida

A disaggregate or microanalytic demand model for urban travel is developed. The model takes into account the simultaneous and interdependent character of decisions about travel, location, and automobile ownership. Two special features characterize the structure of the model: The household is taken as the basic decision unit, and the social class and the stage of the life cycle of those units are important explanatory variables. Stratifying Boston survey data according to those groups generally supported the hypotheses and emphasized the importance of location as a prior determinant of travel choices. An important conclusion for transportation policy grows from that observation: Indiscriminant improvements in transit service, which do not consider the existence of market segments defined by location, may lead to frustratingly small changes in the use of public transportation.

•MUCH EFFORT has been allocated to the design of models for predicting future demands for transportation. Through the contributions of hundreds of individuals, the emphasis has gradually shifted from a purely pragmatic interest in the forecasting of volumes of travel to a more fundamental concern with the explanation of the underlying factors that determine the response of the population when confronted with transportation choices. This paper describes recent efforts to develop such a causal model for urban travel.

A disaggregate or microanalytic demand model for urban travel is developed. It takes into account the simultaneous and interdependent character of decisions about travel, location, and automobile ownership. Two special features characterize the structure of the model: The household is taken as the basic decision unit, and the social class and the stage of the life cycle of those units are proposed as important explanatory variables. As shown below, the results generally support the hypotheses and emphasize the significance of location as a prior determinant of travel choices.

LEVEL OF ANALYSIS

Most of the existing urban transportation demand models deal with the behavior of aggregate masses of population, such as those residing in geographical zones of a city (39). That may not be appropriate. First, it is not clear how the choice of residential location can be accounted for as an explanatory factor of demand. It is also rather difficult to identify the effects of transportation characteristics that may be significant to individuals but relatively unimportant in explaining the behavior of zonal populations. Finally, macroanalytic models embody the critical assumption that households within a given zone are fairly homogeneous and that variations in zonal averages accurately reflect the variations among individuals. But as McCarthy (21) and others have shown, this hypothesis does not appear correct for transportation in light of available evidence.

The alternative to an aggregate or macroanalytic model is a disaggregate or microanalytic analysis. Whereas the macroanalytic analysis estimates the parameters of a model from data on the average behavior of groups of the population, the microanalytic

Publication of this paper sponsored by Committee on Social, Economic and Environmental Factors of Transportation.

approach focuses on information about individual units. Statistically speaking, disaggregate analysis provides more efficient estimates of the parameters in that smaller standard errors can be obtained at a smaller computational cost. Furthermore, a disaggregate analysis avoids the ecological fallacy of inference, whereby factors that coincidentally dominate the behavior of the arbitrary groups of an aggregate analysis are interpreted to affect the behavior of individuals (8). On both counts, a microanalytic analysis is preferable for developing demand models.

The interest in microanalytic models can be justified on 3 more counts. First, disaggregation provides a most natural setting for the development of causal relations among their components, based on simple assumptions about the behavior of the decision-making unit. Second, they usually allow a building-block approach that can be extremely useful as a strategy for the development of urban models based on interrelated blocks describing the urban transportation, housing, educational, and other sectors. Finally, they provide useful guidance as to the appropriate way to aggregate data and relations in the development of more efficient and operational aggregate models (32).

Recently, there has been an increasing interest in a more disaggregate analysis of complex social systems (25). In urban transportation, the microanalytic approach has generally been focused on the prediction of the commuter's selection of a transportation mode for his journey to work (19, 22, 27, 31, 34). Most of those models have been able to deal successfully with detailed attributes of the mode of transportation and of the travelers.

CONTEXT AND DECISION UNITS

Many of the microanalytic transportation demand models treat as exogenous important characteristics of the traveler, such as whether an automobile is available for the journey to work and whether transit constitutes a valid alternative. But those attributes are actually the results of choices that the decision-maker or his household has made, within a broader system, to satisfy demands for accessibility to sources of income, commodities, and services. It is natural, therefore, to extend the microanalytic model to this larger context. By so doing, the focus is switched from the modal-split problem to the demand for urban transportation and from the trip to the different means by which the decision unit can meet its demand for mobility.

As a further extension, the decision unit, critical to any microanalytic model, was chosen in this study as the household rather than the individual. Given the present structure of society, mobility and travel decisions are, by and large, made implicitly or explicitly by the household. That selection is both intuitively reasonable (5) and well supported by considerable evidence. Oi and Shuldiner (23) showed how much understanding of urban travel could be gained by examining individual households, and researchers in the closely related area of demand for durable goods have learned to explain demand through the study of family budgets and characteristics (26). That experience is extremely relevant to urban transportation demand because of the dominance of the automobile, one of the most important durables, in urban travel.

MODELING CONSIDERATIONS

To model the demand of households for mobility in an urban environment is to achieve a representation of the outcomes of the process by which each household selects sets of transportation options under given sets of stimuli. Ideally, one would like to achieve a dynamic representation that could "simulate" the adjustment, through time, of each household to different stimuli. That approach requires a clear understanding of, and detailed information about, the process underlying the choices and behavior of households at any moment. As a prerequisite for gaining that understanding, one must usually begin by learning how the household's course of action is affected by different situations.

The concept of equilibrium, when applicable, has been found to be extremely useful for this purpose. Equilibrium assumes that any random observation on a system shows it at, or very close to, the most stable or desirable position, given the set of stimuli upon it. For example, in the urban transportation system, a household might be assumed to have the most preferred number of cars in accordance with any other alternatives, its socioeconomic characteristics, and whatever constraints may limit its choices. To the extent that a system is, indeed, close to equilibrium, the governing models are much easier to estimate than they would be for a dynamic model.

The simplification obtained by introducing the concept of equilibrium is paid for. It imposes caution on the interpretation that can be given to the response of households to external factors. Specifically, Grunfeld (12) and Malinvaud (20) have shown that coefficients estimated from cross-sectional data tend to include long-term tendencies and cannot, by themselves, provide fully accurate estimates of short-run responses. But that limitation does not seem to hamper the exploration of the causes of individual choice.

Acceptance of the notion of equilibrium points to a set of postulates drawn from the theory of consumer behavior, as described by Lancaster (16). This theory, which has already provided a basis for much work in demand studies, can be used as a theoretical base for the specific model formulated here. The specific postulates are that households (a) desire transportation characteristics such as mobility and accessibility, which are required in their daily activities and (b) attempt to maximize, subject to the constraints of available income, the combined utility that they can obtain from the characteristics of all services and commodities, including mobility and accessibility.

CHOICE OF VARIABLES

To specify the model, it is necessary to be fully explicit about its endogenous and exogenous variables. These are described below from a theoretical point of view. The translation of such definitions into practical terms always requires some considerable effort. In this instance, data from 3 sources were used for the application of the model to Boston: the comprehensive traffic and transportation inventories (38), which include files on more than 38,000 households, 117,000 persons, and 300,000 trips; the land use and forecasting matrices for 626 zones, which were developed by the Eastern Massachusetts Regional Planning Project; and statistics of the Registry of Motor Vehicles. Details on the use of those files for the establishment of the variables are given by Aldana (1).

Endogenous Variables

The endogenous variables of this model can be considered to be the dimensions of the space in which a household can look for sources of mobility and accessibility. Neglecting some rather unusual cases, there are essentially 3 dimensions: residential location, number of automobiles available to a household, and use of public transportation.

The choice of residential location is, certainly in part, a transportation decision, notwithstanding evidence that households are often concerned more with neighborhood quality than with accessibility in choosing a site (3, 30). Location clearly influences their available options and, thus, their choices. Conversely, households that can select some options (for example, buy several cars) are more likely to locate in certain zones than those that cannot. Although metropolitan areas are quite heterogeneous, 2 broad subareas seem to be especially important from the point of view of accessibility and mobility: the business districts with their concentrated activities and the suburban residential area. The household's choice of location is, subsequently, regarded as a selection of one of those two.

The number of automobiles available to a household (say, none, one, or more) is both a determinant of choice of transportation means and, as a consequence of residential location, a result of preferences for different forces of mobility. That levels of automobile ownership are determined outside of the model has been an assumption in most empirical studies, but it seems more reasonable to take this as a transportation decision. A few recent studies by Kain (14), Shindler and Ferreri (29), and Leathers (18) have already taken that view.

The use of public transportation would be the central issue for the applications concerned with the prediction of the effects of improvements in this sector. A natural classification of the households would seem to be between those using transit regularly and the rest. This typology indicates not only which households rely in some sense on transit as a source of mobility but also which households are likely to have information about transit schedules and transit times.

Exogenous Variables

As in any other model, the exogenous variables can be classified into policy and control variables. The policy variables are those that can be altered to achieve particular objectives and, thus, whose effect the analyst wishes to predict. In the present model, they are represented by variables measuring the level of service of public transportation.

The control variables measure the diversity of the population being observed. In the present case, they are the measures of the socioeconomic characteristics of households. To avoid errors in the estimation of the effects of the policy variables, one should stratify or segment the population and the data into reasonably homogeneous groups, as described by the control variables, or those latter variables should be explicitly included in the model.

Market Segmentation of Households

The stratification of the population of households into groups likely to have similar utility functions is a prerequisite for a microanalytic model based on the analysis of the behavior of such micro-units. This need has been recognized in the past from quite different points of view. Thus, Zellner (41) studied the biases introduced into the analysis when the micro-units are different, and marketing researchers have been concerned with this problem under the heading of "market segmentation" (11). Both of those points of view are complementary. The stratification of the sample population is necessary to ensure the statistical acceptability of the estimates and, once those estimates are found, permits an analysis of the differential characteristics of the population strata.

The study of consumer behavior has shown that it is extremely important to consider the so-called "life cycle stage" of the household, which accounts for a large fraction of the variation of consumptional patterns of the households, and of automobile ownership in particular (15). Lansing and Morgan (17), for instance, suggest 3 main stages in the life of an ordinary family—the bachelor stage, the stage of marriage, and the stage of the solitary survivor—and describe how income, expenditures on durable goods, and attitudes about financial position differ from stage to stage. The closely related concept of age has, of course, been extensively used as an explanatory factor in studies of urban transportation demand (37, 40); but, as indicated by Wells and Gubar (36), the life-cycle concept seems to provide a better description of the family as a unit.

Another important taxonomic concept is what sociologists have referred to as social class. Although there is no consensus on the definition of that term, a social class may generally be thought of as a group of individuals with broadly similar positions of power. Some marketing researchers, such as Carman (4), have found the concept highly useful. Specific justification for the use of social class in the study of the demand for urban transportation comes in addition from empirical studies of urban social stratification, which disclose a close relation between social and spatial distances in a community (10), and from evidence of trip generation rates of different occupational groups (33).

Income was not used for the segmentation of households. First, several researchers have, as reported by Carman (4), become convinced that social class is a more significant determinant of consumption patterns. Second, the residual effects of income, once social class and life-cycle stage have been considered, can relatively easily be approximated.

Classification of Data

The data for the example analysis in Boston were classified quite readily into the postulated typologies of location, life cycle, and social class. Standard clustering techniques, such as factor analysis, were used (13). The essential procedure consists of using detailed data on each element of a population to define mutually exclusive, collectively exhaustive groups. A formula, which is developed by using any one of a number of procedures, assigns elements to the groups so as to minimize the chance of misclassification. Details of the techniques used are given by Aldana (1).

For the example analysis, all Boston zones were divided into the 2 categories of business district and suburban. The following market segments were obtained. Seven life-cycle stages were significant: young bachelors; childless, young couples; couples with small children; couples with teenagers or adult dependents; broken families; childless, old couples; and single, old persons. Only 2 social classes were significant: white- and blue-collar workers.

THE CHOICE MODEL

The description of the model can be made more precise if possible outcomes of the choice are regarded as alternative "states" in a 3-dimensional space. From the point of view of the household, the states are described by combinations of the 3 endogenous variables: automobile ownership level, transit usage, and locations within the metropolitan area. The choice procedure can be simply regarded as the activity of the household directed toward evaluating the utility of, or preference for, each one of those states and the selection of the one affording it the highest satisfaction.

Because the model does not attempt to explain the choice of residential location, the problem arises as to how to consider that aspect. Specifically, there is the possibility that households that select locations of high business and economic activity, as opposed to more residential locations, do so because they have different utility functions from the others. To accommodate that likely contingency, the sample population was stratified so that each household's preferences could depend both on the location of the household's residence and on the 14 demographic groups (7 life cycles times 2 social classes).

The preferences of each household were defined in terms of their utility. That is a measure of their value of any alternative and can only be expressed in relative units. (Technically, the utility is measured on an ordered metric scale, constant up to a positive linear transformation.) For simplicity, the utility was taken to be linear. In symbols, the utility for household i of the transportation option kl at a given location j may be expressed as

$$U^{i}_{kl/j} = d_{kl\,o} + d_{kl\,p}X_{ij} + e^{i}_{jkl} \tag{1}$$

where $d_{kl\,o}$ is a constant, $d_{kl\,p}$ is a vector of parameters, X_{ij} is a vector of characteristics of the household i and the transportation options at location j, and e^{i}_{jkl} is the error term or disturbances representing omitted characteristics and elements not accounted for explicitly by the model.

The basic hypothesis concerning the choice procedure is that preferred alternatives are chosen. This requires that

$$U^{i}_{k*l*/j} \geq U^{i}_{kl/j} \tag{2}$$

for all i, k, l, where k*l* is the transportation option selected by household i, and the subscript j indicates that the outcome is influenced by the location of the residence of the household.

At this point, it would be possible to make suitable assumptions about the stochastic characteristics of the disturbance and to devise a method for estimating the parameters of Eq. 1 by using as a criterion, for example, minimization of the number of misclassifications in the sample. It seems convenient, however, to examine in greater detail the state at which each household is observed in order to further illustrate the scope of the model as well as the extent of its assumptions.

The transportation choices open to the households are the distinct, alternative states described by the endogenous variables. In this case, there are thus 12 discrete points: 2 possible household locations times 3 levels of automobile ownership and 2 levels of transit usage. Those can be represented by a vector S, whose components s_j have the following properties:

$$s_j = 0 \text{ or } 1 \text{ for } j = 1, 2, \ldots, 12$$
$$\Sigma_j s_j = 3 \tag{3}$$

Suppose that it is possible to identify and to measure for every household a vector of variables T, which can be assumed to be causally linked to S. In a comparison with the formulation of Eq. 1, the vector T includes those variables in X_{1j} plus the variables affecting the locational choice, and the link between T and S is provided by relations similar to Eqs. 1 and 2. It should be clear that there is strong interaction among the 3 main components of S (i.e., location, automobile ownership, and transit usage affect one another).

From the mathematical and statistical points of view, the situation is clearly one of a system of simultaneous relations where there are some exogenous variables represented by T and some endogenous variables represented by the components of the vector S. The values taken by any of the components of S are jointly influenced by T and by the values taken by the other components of S.

Although the system described above is not exactly the kind of system of simultaneous relations found in econometrics (6, 20), much insight is gained by comparing it with such systems. First, they are different in that in the system presented here the endogenous variables are discrete in nature and are not related, among them and with the exogenous variables, by simple linear relations as is the case in most systems dealt with in econometric applications. On the other hand, statements about the conditional distribution of one or more of the components of S, given the other components of S and the vector T, are the analog of what econometricians denote by the "structural form" of the system, that is, the set of relations that are assumed to be autonomous and stable and those on which the model builder can impose restrictions derived from his knowledge about the behavior of the system. On the same line of thought, the marginal or unconditional statements about the distributions of components of S, independently of any other component of S, are the equivalent of the so-called "reduced form" of econometric systems.

That comparison of the model with econometric systems of simultaneous equations provides intuitive but rational arguments to develop, by analogy, large sample techniques of estimating the parameters of the model. As in econometrics, the main focus of interest is in the "structural form" of the systems. Its parameters can be readily estimated by appropriate stratification of the sample, as explained below.

Continuing with the choice model in Eq. 3, let S_1 be the subset of components of S related to location, S_2 be the subset of components of S related to automobile ownership and level of transit usage, S_{2i} be any specific transportation option, and $P(S_{2i}/S_1T)$ be the conditional probability that a household will select option S_{2i} given the choice of location and the vector of characteristics T.

The distribution of the vector T given S can be regarded as the joint distribution of the variables making up the vector T in each one of the "cells" determined by S. Standard assumptions are that the variables in T (or a suitable transformation of them) are jointly normally distributed with a common variance-covariance matrix; each one of the cells is determined by S_2. Under that type of assumption and by the use of Bayes' rule, it is easy to show that statements such as

$$P(S_{2i}/S_1T) = \frac{q_{2i} P(T/S_{2i}S_1)}{\Sigma_j q_{2j} P(T/S_{2j}S_1)} \tag{4}$$

lead to the probability of group membership in the standard multigroup discriminant analysis (2, 24, 28, 35). It could also be possible to relax somewhat the assumptions made and use alternative techniques such as the multinomial extension of logit analysis (1). Therefore, it is possible to estimate the coefficients of linear functions similar to Eq. 1, which discriminate among the several transportation options, by conditioning (stratifying the samples) on the choice of residential location. Those discriminant functions, after due imposition of constraints in the coefficients of the variables that do not affect the choice of the transportation option, can in fact be regarded as the relative utility function specified in Eq. 1.

If the purpose of the analysis is to quantify the effects of marginal changes in the policy variables, as when elasticities are computed in economic demand studies, the work is finished once the structural forms have been estimated. Moreover, if the model is to be used in making conditional predictions of one of the endogenous components, and it is realistic to assume that the rest remain constant or are known, then the structural form is all that is required. However, in the general prediction case it is necessary to have the reduced form. These unconditional statements can be efficiently obtained from a combination of the conditional statements about residential location, automobile ownership level, and transit usage, singly or appropriately combined (1).

ILLUSTRATION OF RESULTS

The Mobility Model

The estimation of the parameters of the whole model is a rather lengthy undertaking, the reason being that this microanalytic approach requires quantities of detail. All elements of the model were, thus, not estimated. Typical results are now presented by means of an illustration. It refers to those parameters related to the transit choices of demographic group 3, white-collar couples with small children, and provides some insight into the general procedures that must be followed in the estimation of the complete model.

Table 1 gives a cross classification of automobile ownership levels versus levels of transit usage for households of group 3 located in the central city and suburban zones. It is quite clear that there are strong interactions among the options open to the households. Therefore, even if one intends to explain the transportation choices only, automobile ownership and transit usage, those should be considered as simultaneous choices, conditional on the locational decision.

An obvious procedure for estimating the parameters would consist of finding the discriminant functions for the 6 possible choices in each location. However, that procedure would require one to consider all the variables involved in the choice of either automobile ownership level or transit usage level, which is the analog of estimating the reduced form in the analysis of simultaneous equations.

The alternate procedure suggested here is the equivalent of 2-stage least squares in econometrics. This allows a more efficient use of prior information about those variables that do not affect each conditional choice. This procedure is illustrated for the choice of transit usage level.

For the first step, we let $P(B_i/L_j A_k X)$ be the probability for a household to use or not use transit (B_i = 1 or 0 respectively) given that it is located at L_j (central city or suburbia), has a level of automobile ownership A_k(0, 1, 2, or more), and is described by the vector of exogenous variables X. These probabilities can be estimated by making use of the 2-group discriminant technique at each level of automobile ownership at each location. It is also possible to impose constraints in the components of X such as excluding those variables that are considered a priori as not affecting the choice. Thus, for example, one might argue that, given that a household in demographic group 3 resides in suburbia and has 2 or more cars, the distance, within limits, to the transit station is of no importance in the transit choice because any of the 2 adult members can drive and park the car at this station. It should be clear that, because one is estimating the probability of one choice given other endogenous choices, one is estimating the structural or conditional form for transit choice.

The coefficients of the discriminant function obtained for those households residing in the suburbia and having 1 automobile are given in Table 2. More than 80 percent of the household choices were correctly classified by this discriminant analysis.

Although it is not appropriate to discuss in detail the conclusions implied by data given in Table 2, a few comments are in order. First, relative transit accessibility, defined as the number of jobs that could be reached within a 15-min ride by transit divided by the number that could be reached within a 15-min ride by automobile, is a key policy variable because it can be affected by changes in the transportation system. Second, walking time to the transit station did not appear to be significant mainly because it was not possible to measure that variable in an appropriate manner. Finally, the signs of the coefficients agree with a priori notions of causality, and the various calculated statistics allow one to reject the null hypothesis about the joint significance of the exogenous variables.

Having this discriminant function, one can readily compute the probabilities of transit choice conditional on location and automobile ownership. The problem now is to estimate the probabilities of transit choice not conditional on the automobile ownership level. To continue with the procedure, let $P(A_k/L_jX)$ be the probability for a household to choose 0, 1, or > 1 automobiles, given that it is located at L_j and is described by the vector of exogenous variables X. These probabilities can be computed from the discriminant functions obtained by considering the 3 automobile-ownership groups at each of the 2 locations, disregarding the transit choice, and including in X all the variables that affect the choice of automobile ownership and transit usage levels. The similarity of this step with the first stage in 2-stage least squares should be noted.

One can now compute the joint probabilities of using or not using transit and having 0, 1, or >1 cars by simple multiplication. Adding these joint probabilities to the levels of automobile ownership, one finds the probabilities of using or not using transit conditional on location but unconditional on the choice of automobile ownership level. This procedure makes it possible to compute the probabilities of transit choice both conditional and unconditional on the levels of automobile ownership.

The models developed were tested by using them to predict the choices made by households outside Boston's circumferential Route 128; the households were not used in the sample from which the coefficients of the model were derived. The results are given in Table 3. The data illustrate the use of increasing levels of information in predicting the percentage of households and the ability of the model to predict the transit usage of households under quite different conditions.

The Trip Model

To predict the number of trips, one should make some assumptions about the direction of causality among the number of trips and the location and transportation options. If the number of observed trips is regarded as affecting the choice of mobility state, one would have to deal with simultaneous estimation procedures and resort to the estimation of 12 different relations, one corresponding to each location and transportation option. But it seems that it is the number of trips "desired" by the household that influences the choice of state. Furthermore, this number is likely to be different from the number of trips actually made in a random day. One might, thus, feel intuitively satisfied that the causal direction is unidirectional, from location-transportation option to number of trips.

Because there are no guidelines for the choice of the specific form of the model for predicting trips, it was decided to use a linear additive form of the explanatory variables, for the sake of simplicity and as a first approximation. Thus, the analytical form of the trip model was taken to be

$$T_d^{pm} = M_d^{pm} + \sum_{jkl} A_{djkl}^{pm} + \underline{B}_d^{pm}\, \underline{X} + E_d^{pm} \tag{5}$$

where T_d^{pm} represents trips made by households in demographic group d for purposes p and by mode m; M_d^{pm} is the mean number of those trips per household; A_{djkl}^{pm} is the additive effect for those trips resulting from the household having chosen location j (central city or suburbia), automobile ownership level k (0, 1, or > 1 automobiles), and level of transit usage 1 (no transit or some transit); $\underline{B}_d^{pm}$ is the vector of coefficients for the exogenous variables; $\underline{X}$ is the vector of exogenous variables; and E_d^{pm} is the error term.

Table 4 gives the estimated coefficients of the covariates, and Table 5 gives the adjusted mean trip rates in each location-transportation option for the trips to work or school by automobile taken by households in demographic group 3. These figures were estimated by standard analysis of covariance techniques. Generally, the estimates are significant, and their magnitudes are in agreement with what common knowledge would have indicated: The more cars and the less transit available, the more automobile trips will be taken.

Table 1. Households in demographic group 3 by location and transportation options.

Location	Automobile Availability	Number of Households: No Transit	Some Transit	Total
Central city	0	13	49	62
	1	175	129	304
	>1	13	5	18
	Total	201	183	384
Suburbia	0	18	26	44[b]
	1	917[a]	285[a]	1,202[b]
	>1	525	60	585[b]
	Total	1,460	371	1,831

[a] Households used to estimate discriminant function in choice of level of transit usage, conditional on level of automobile ownership.

[b] Households used to estimate discriminant function in choice of automobile ownership level, unconditional on transit choice.

Table 2. Coefficients of discriminant function for choice of transit for suburban household with 1 automobile.

Variable	Coefficient	Significance Level (percent)
Spouse working (1 yes, 0 no)	0.903	1
Children under 5 years of age (0 yes, 1 no)	0.527	1
Working outside of residential zone (1 yes, 0 no)	0.400	
Working in central district (1 yes, 0 no)	3.124	1
Household income, thousands	0.0135	5
Relative transit accessibility (transit/auto)	1.992	1
Walking time to transit station, min	-0.0203	

Note: F-value, 51.6; degrees of freedom, 7 and 1,194.

Table 3. Percentage of transit use explained by model for households outside Route 128.

Forecasting Method	Predicted Use	Difference From Actual	Fraction Explained
Actual transit usage	6.2	—	—
Households using transit inside Route 128	39.3	33.1	—
Households in demographic group 3 using transit inside Route 128	25.0	18.8	43.2
Suburban households in demographic group 3 using transit inside Route 128	20.2	14.0	57.5
Full model prediction	8.9	2.7	92.0

Table 4. Estimated coefficients of significant exogenous variables for trips to work or school by automobile for demographic group 3.

Variable	Coefficient	Standard Error
Household size in number of persons over 5 years of age	0.20	0.05
Spouse working (1 yes, 0 no)	1.16	0.22
Participation in car pool (1 yes, 0 no)	2.35	0.66

Note: Sample size = 2,200 households.

Table 5. Estimated adjusted means of trips to work or school by automobile for each transportation option in demographic group 3.

Location	Automobile Availability	No Transit: Adjusted Means	No Transit: Standard Error	Some Transit: Adjusted Means	Some Transit: Standard Error
Central city	0	1.51	0.83	0.49	0.43
	1	2.90	0.23	2.01	0.27
	>1	3.70	0.83	4.82	1.35
Suburbia	0	1.16	0.73	0.42	0.59
	1	3.14	0.10	2.05	0.18
	>1	3.98	0.13	2.61	0.39

Note: F-value, 15.0; degrees of freedom, 11 and 2,185.

CONCLUSIONS

The model produces results that seem admissible in the light of prior knowledge of the situation, and tests of hypotheses rejected the irrelevance of the postulated causal mechanism. Statistical tests of the validity of the model, although never conclusive, did not undermine the credibility of its predictions.

The choice and trip models provide a means to assess the impacts of policy changes on the different segments of society. Concepts borrowed from the social sciences, such as life-cycle stage and social class, seem to be extremely useful for segmenting the population into homogeneous strata, a necessary step in any disaggregative approach. Only through the identification of the several market segments and the quantification of the intensities of their responses will it be possible to design truly effective strategies for expanding the clientele of urban transit.

One result of the limited calculations conducted so far is the indication of low sensitivity of transit usage for a particular demographic group to changes in accessibility. This contrasts with the large variations shown by this group throughout the entire city. One is led to conclude, therefore, that the changes in use of transit observed within the city are not just the result of differences in the level of transit service but are mostly the product of dissimilar attributes and preferences of the households located in different areas. As a corollary, it would appear that indiscriminate improvements in transit service, that is, changes that do not consider the existence of market segments, might lead to frustratingly small changes in transit use, at least until long-term changes in residential patterns had adjusted to existence of new service.

ACKNOWLEDGMENTS

The authors are grateful to many groups for the support, both intellectual and financial, that encouraged this work and enabled it to be completed: the Ford Foundation, the M.I.T. Department of Civil Engineering, and the M.I.T. Urban Systems Laboratory.

REFERENCES

1. Aldana, E. Toward Microanalytic Models of Urban Transportation Demand. Dept. of Civil Eng., M.I.T., Cambridge, PhD thesis, 1971.
2. Anderson, T. W. An Introduction to Multivariate Statistical Analysis. John Wiley, New York, 1958.
3. Butler, E. W., Chapin, F. S., Hemmes, G. C., Kaiser, E. J., Stegman, M. A., and Weiss, S. F. Moving Behavior and Residential Choice—A National Survey. NCHRP Rept. 81, 1969.
4. Carman, J. M. The Application of Social Class in Market Segmentation. Inst. of Bus. and Econ., Univ. of California, Berkeley, 1965.
5. Chapin, F. S. Explaining Consumer Behavior On What Level. In Consumer Behavior (Foote, N.N., ed.), New York Univ. Press, New York, Vol. 4, 1961.
6. Christ, C. F. Econometric Models and Methods. John Wiley, New York, 1966.
7. Cragg, J. G. Programs for Multiple Probit and Logit Analysis and Extensions to Them. Unpublished, 1968.
8. de Neufville, R., and Stafford, J. H. Systems Analysis for Engineers and Managers. McGraw-Hill, New York, 1971.
9. Development of a Behavioral Urban Travel Demand Model. Charles River Assoc., final report draft, June 1971.
10. Duncan, O. D., and Duncan, B. Residential Distribution and Occupational Stratification. Am. Jour. of Sociol., Vol. 60, 1954-1955, pp. 493-502.
11. Frank, R. E. Market Segmentation Research: Findings and Implications. In Applications of the Sciences in Marketing Management (Bass, King, and Pessemier, eds.), John Wiley, New York, 1968.
12. Grunfeld, Y. The Interpretation of Cross Section Estimates in a Dynamic Model. Econometrica, Vol. 29, July 1961, pp. 397-404.
13. Harman, H. M. Modern Factor Analysis. Univ. of Chicago Press, 1967.

14. Kain, J. F. A Contribution to the Urban Transportation Debate: An Econometric Model of Urban Residential and Travel Behavior. Rev. of Econ. and Stat., Vol. 46, 1964, pp. 55-64.
15. Katona, G., Mueller, E. L., Schmiedeskamp, J. W., and Sonquist, J. A. Survey of Consumer Finances, 1966. Inst. for Social Res., Univ. of Michigan, Ann Arbor, 1967.
16. Lancaster, K. J. A New Approach to Consumer Theory. Jour. of Polit. Econ., Vol. 74, April 1966, pp. 132-157.
17. Lansing, J. B., and Morgan, J. N. Consumer Finances Over the Life Cycle. In Consumer Behavior (Clark, L. H., ed.), New York Univ. Press, New York, Vol. 2, 1955.
18. Leathers, N. J. Residential Location and Mode of Transportation to Work: A Model of Choice. Transp. Res., Vol. 1, 1967, pp. 129-155.
19. Lisco, T. E. Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, PhD thesis, 1967.
20. Malinvaud, E. Statistical Methods of Econometrics. Rand McNally, Chicago, 1966.
21. McCarthy, G. M. Multiple Regression Analysis of Household Trip Generation—A Critique. Highway Research Record 297, 1969, pp. 31-43.
22. McGillivray, G. R. Binary Choice of Transport Mode in the San Francisco Bay Area. Dept. of Econ., Univ. of California, Berkeley, PhD thesis, 1969.
23. Oi, W. Y., and Shuldiner, P. W. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, Ill., 1962.
24. Olkin, I., and Tate, R. F. Multivariate Correlation Models With Mixed Discrete and Continuous Variables. Ann. of Math. Stat., Vol. 32, 1961, pp. 448-465.
25. Orcutt, G. H., Grennberger, M., Korbel, J., and Rivlin, A. M. Microanalysis of Socio-Economic Systems: A Simulation Study. Harper and Row, New York, 1961.
26. Pyatt, F. G. Priority Patterns and the Demand for Household Durable Goods. Univ. of Cambridge Press, England, 1964.
27. Quarmby, D. A. Choice of Travel Mode of the Journey to Work: Some Findings. Jour. of Transp. Econ. and Policy, Sept. 1967, pp. 273-314.
28. Rao, C. R. Linear Statistical Inference and Its Applications. John Wiley, New York, 1965.
29. Shindler, R., and Ferreri, M. G. Auto Ownership as Affected by Transportation System Alternatives. Traffic Eng., Vol. 38, Oct. 1967, pp. 24-28.
30. Stegman, M. A. Accessibility Models and Residential Location. Jour. of Am. Inst. of Plann., Vol. 35, Jan. 1969, pp. 22-29.
31. Stopher, P. R. A Probability Model of Travel Mode Choice for the Work Journey. Highway Research Record 283, 1969, pp. 57-65.
32. Stowers, J. R. Residential Simulation Models for Urban Planning: A Methodological Study With Emphasis on Taxonomy and Location Decision Rules. Northwestern Univ., Evanston, Ill., PhD thesis, 1968.
33. Walker, J. R. Social Status of Head of Household and Trip Generation From Home. Highway Research Record 114, 1966, pp. 141-151.
34. Warner, S. L. Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice. Northwestern Univ. Press, Evanston, Ill., 1962.
35. Warner, S. L. Multivariate Regression of Dummy Variates Under Normality Assumptions. Jour. of the Am. Stat. Assn., Vol. 58, 1963, pp. 1055-1063.
36. Wells, D., and Gubar, G. Life Cycle Concept in Marketing Research. Jour. of Mark. Res., Vol. 3, Nov. 1966, pp. 355-363.
37. Wickstrom, G. V. The Use of Census Data in Urban Transit Planning. Unpublished, 1970.
38. Comprehensive Traffic and Transportation Inventory. Wilbur Smith and Assoc., 1965.
39. Wohl, M., and Martin, B. V. Traffic System Analysis for Engineers and Planners. McGraw-Hill, New York, 1967.
40. Wynn, F. H., and Levinson, H. S. Some Considerations in Appraising Bus Transit Potentials. Highway Research Record 197, 1967, pp. 1-24.
41. Zellner, A. Estimation of Cross-Section Relations: Analysis of a Common Specification Error. Metroeconomica, Vol. 14, 1962, pp. 11-117.

A MULTIMODAL LOGIT MODEL OF MODAL SPLIT FOR A SHORT ACCESS TRIP

Paul Inglis, De Leuw Cather and Company of Canada, Ltd.

The paper involves a discussion of research into the use of a logit mathematical formulation to model modal split. The model investigates the extension to the multimodal situation in terms of both user and system variables. Time and cost difference are the major variables introduced for the systems; income, age, and a rush-hour dummy are other variables entered. The trip modeled is a short commuter trip, which is a part of a longer overall trip and involves the access to a commuter rail station only. The line-haul portion is not considered. The mathematical formulation is probabilistic in nature and can be used with any number of choices of mode. Four choices were available for the research data. Two methods of aggregating, summing probabilities or taking absolute choices, were discussed and tested. The value of time can be developed by the use of time and cost coefficients derived for the model. A value was found that was similar to some of the previous values but of lesser magnitude than values found for longer trips. The model coefficients themselves were derived by a computer program that employed a maximum likelihood technique to iterate to significant values.

•STUDIES to improve the flow and direction of traffic volumes have been conducted for many years. As early as 1844, traffic counts were being made in France. Yet it was not until federal legislation in the United States in 1944 that transportation planning, in the form of origin-destination studies, evolved in a recognizable form. Only since 1955 has there been any advance beyond simple extrapolation of past trends. Modern analytic predictive planning, then, has been developing for only a relatively short span of less than 20 years. From the relevant technology, there has evolved a relatively standard format for predicting future flows. This has been called the "urban transportation planning" (UTP) package.

The package generally comprises 4 models: trip generation, trip distribution, modal split, and network assignment. Martin, Memmott, and Bone (15) and Davis (5) discuss the UTP package in full but cannot agree on a precise sequence for the 4 models. The first authors prescribe modal split to remove the transit riders before distribution and assignment, while the latter inserts modal split after distribution to try to achieve a total picture.

This paper concerns only one part of the UTP package: the choice of mode for a relatively short journey.

Since the late fifties, major transportation studies have been carried out in almost all major cities in North America. Each study was required to build its own models for future prediction. A definite advantage can be gained if some degree of standardization can be obtained. That has not previously been possible because of the nature of the explanatory variables employed in the models. Early work by Wynn (25), Carroll (4), and Adams (1) placed a large emphasis on urban land use and zoning. That was followed by Chicago Area Transportation Study reports by Howe (12), Biciunnas (3), and Sharkey (20, 21) and a Milwaukee study report by Hadden (9). They placed emphasis only on socioeconomic measures and activity levels derived from urban zonal theory.

Publication of this paper sponsored by Committee on Social, Economic and Environmental Factors of Transportation.

This meant that changes in the systems would not be reflected in the results of the models. A second shortcoming in the use of variables such as income and car ownership was the continual inflationary trend as the standard of living increased. Both of the above groups used linear techniques in the model formulas and aggregate groupings in zones as the basic unit for prediction.

The use of high levels of aggregation resulted in a very high variance within the zones, especially when short trips were being considered. To illustrate, a trip from zone A to adjoining zone B could vary from several blocks to more than a mile. Also, generalized zonal activity could override important small pockets of different types of land use.

By the late fifties, interest in system variables had begun to rise. Large studies in Washington, Chicago, San Francisco, Toronto, and Philadelphia resulted in the definition of a set of diversion curves for modal-split prediction. The models related either time differences or time ratios to the percentage of total trips diverted from one mode to the other. The techniques used are well documented by Quinby (18), Hamburg and Guinn (10), and Hill and Von Cube (11). Quinby recognized the mathematical inferiority of regression in this case and proposed that Pearl-Reed logistic curves be fitted before he finally settled on a Gompertz exponential curve formulation. The above works led to a large set of diversion curves illustrating the diversion to transit with a change in travel time ratios, for different cost ratios and income levels. They were developed by Traffic Research corporation and documented by Deen, Mertz, and Irwin (7). However, the portions of the curves of highest predictive value were also the areas of greatest uncertainty, for no corroborating observations were available for the predictive areas.

Errors in those earlier models were potentially very high. Much of the error in prediction was attributed to the level of aggregation at which the models were built. Working at the zonal level did not allow the large variance within the zonal populations to be accounted for. Reducing to the basic component, the individual user, overcomes that problem. The model can then be aggregated to any level desired with less chance of variance errors occurring.

Modal split provides a good point to start from in redeveloping the UTP package at the disaggregate level. Data are relatively easy to gather, the result can be easily measured by survey, and the variables of influence can be defined.

Once the modal-split model has been completed, it is anticipated that the use of a similar technique and mix of user and system characteristics will allow the whole package to be integrated. Several authors who have done work with stochastic models of modal split include Warner (25), Quarmby (17), and Lisco (14). Currently 3 separate classes of models exist, depending on the type of statistical technique: discriminant analysis, probit analysis, and logit analysis. The use of linear regression has been largely superseded because of the possibility of predicting negative probabilities and values greater than one.

Discriminant analysis (8) is based on the existence of overlapping normal subpopulations that are distinct in the decision sense. By identifying attributes that can account for the difference in choice, a function that discriminates among the populations can be developed. Models by Quarmby (17) and McGillivray (16) employed user characteristics as well as system characteristics, but the set of variables and their form are still a subject for much debate and research. Probit analysis was first suggested for modal split by Warner (25), who rejected it as computationally too complex. Lisco (14) was the first to use this method successfully for his economic studies on the value of time, a result derived from the cost and time coefficients of his modal-choice model. Lave (13) has since built other modal-split models by using this mathematical form. It is a technique that requires a normal distribution of threshold values, which may not always be a valid assumption.

Logit analysis was developed by Thiel (24) in the multinomial sense. Stopher (23) limited his use of this technique to binary models in his research at Northwestern University. Rassam, Ellis, and Bennett then developed a multimodal aggregate model of modal split for access to a Washington airport.

This paper attemps to document the next step: the development of a multimodal, stochastic, disaggregate model of modal split for a short journey.

THE MODEL

In the past, models of modal split have usually been restricted to 2 dimensions. When more than a binary choice exists, a set of binary models is required to produce the final result. The design of this set of models requires that some step-by-step choice process be presumed among the modes. If this were not the case, a great number of models would result, for each mode would be compared individually to each of the alternate modes. It is then desirable to be able to compare all modes at once, without having to make an a priori decision on which modes are to be directly compared. The substitution of an n-dimensional model to replace the binary system should be a progressive step. A brief outline is given below of the form of the logit model.

The choice of a modal-split model is based on the premise that the probability of using a particular mode is a continual function whose dependent variable p ranges from 0 to 1 according to some function of the sociological traits of the user and the characteristics of the modes involved. Thus, as any of these variables change, so does the value of the function and hence the probability. The use of a simple linear relation is rejected because of the bounds imposed by the 0-to-1 range. The function should be asymptotic to both of those limits. This can be done using a logit formulation. The binary logit relation may be generalized as

$$p = \frac{e^{G(x)}}{1 + e^{G(x)}}$$

$$(q = 1 - p)$$

$$q = \frac{1}{1 + e^{G(x)}}$$

where G(x) represents a function to describe the response relation to a particular mode. It may be formulated as

$$G(x) = \text{constant} + \sum_{k=1}^{N1+N2} a_k x_k$$

where N1 is the number of system-dependent variables (such as time or cost) and N2 is the number of system-independent variables (such as age or income).

A simple mathematical manipulation in terms of variable differences illustrates the symmetry of the formula. The system-dependent variables are used in terms of differences to reflect the advantage of one over the other while only a single function is dealt with. That symmetry may be extended to the multidimensional form by proposing

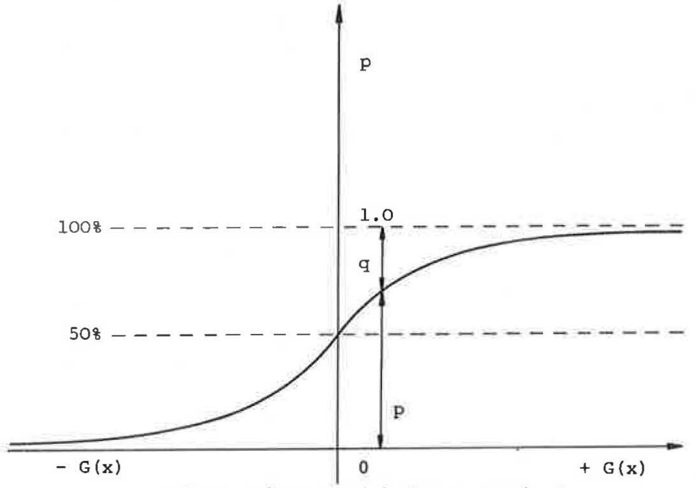
$$p_l = \frac{e^{G_l(x)}}{\sum_{m=1}^{M} e^{G_m(x)}}$$

A pictorial representation is given in Figure 1 of a binary-choice situation.

The major problem, given the relation above, is the estimation of the coefficients in each equation. That was done by the use of a maximum likelihood estimator procedure.

The estimation technique requires that a base mode be established. The generalized formulation above does not lend itself to the maximum likelihood estimator technique. For that reason, the model is developed in terms of (n - 1) different modes, and the last mode is determined by the limitation that the sum of the probabilities is one. To involve the system variables of the last (or base) mode requires that these particular variables be somehow related to the base mode. That can be done by using differences or by using ratios. Both techniques are valid; however, for this model, the former was

Figure 1. Typical 2-dimensional logit curve.



NOTE: Given a G(x) for a particular observation, p and q are the resulting probabilities.

chosen. To have a model that can be realistically analyzed and that conceptually follows the formulation stated above required that the coefficient of the system variable be kept constant for each mode. Also each system variable must be related to each mode. For example, a comfort rating could not be entered into the bus mode if it were not entered in the car mode or in every other mode. On the other hand, the user characteristics need not be entered in each mode equation and should not have like coefficients unless there is a like correlation.

Several advantages are attributable to this technique, particularly in terms of theoretical assumptions. There is no assumption of normality to be met in the G(x) function, resulting in a more generally applicable model. Also the use of a probabilistic sum for aggregation gives a better conceptual idea of the true process that occurs in this mechanism.

To make this model operational requires some aggregation because collection of the information for each trip would be prohibitive. The methodology for this aggregation is still a question for review and testing. There are 2 distinct possibilities. First, if our sample is 10 percent, then each datum is a proxy for 10 other members of the population. It may be assumed that all 10 will make the same absolute choice as the proxy, i.e., the mode of highest probability according to the model. Thus, to aggregate, one multiplies the result by the sample ratio. Second, again if the sample is the same, the probabilities for the individual mode choices and not the ultimate choice are considered. To aggregate, one obtains the sums of the probabilities and multiplies those sums by 10. This better describes the behavioral process because we are dealing with human beings who can and will change their habits in this nonexact manner. The scope of this work included only a minor attempt to determine the superior methodology. That is a matter for further study.

DATA BASE AND VARIABLES

The data used to derive and test the models have been titled the "suburban station access" data. They were collected from people using the Chicago and Northwestern Railroad suburban routes from the northwest corridor into Chicago. The trips that were modeled were not trips to the center of the city but rather shorter access trips between home and the commuter station. The 4 modes involved in the study were walk, drive and park, driven, and bus. A final set of 117 observations was used to build the models with a good balance of each modal user. A second set of 400 observations was used to test the models. The second set had a different modal user mix and slightly different geographical area. Those 2 factors would help in providing a good test of the technique.

Probably the most important task for the model builder is the choice of variables to be used in the models. Below is a brief discussion of the variables that were available from the given data.

Cost

Because cost is the measure of almost all other goods and services, it should be significant in the decision to use a mode of travel. However, true costs are seldom evaluated by the user; rather he sees only direct costs such as parking, gas, and tolls. His decision, then, is based only on this perceived cost. Betak (24) feels that this variable could prove much more important if the costs of the different modes were economically substitutable; however, the investment in a car, for example, is basic to a family and not related directly to a particular trip. That is a system variable and as such is entered as a difference between the mode and a base mode. Because the walk mode has no monetary cost associated with it, it is used as the base mode.

Time

Time is another important system characteristic. Attitudes toward time vary according to the activity during a given period. Therefore, time was entered for waiting, walking, parking, and access time separately; but, except for walking time, no advantage was obtained over total travel time. Overall time, then, is entered as a difference variable in the model.

Time of Day

A rush-hour dummy variable was entered to try to ascertain any difference in attitude between the morning rush period and the rest of the day. The dummy has a value of 1 if the trip is taken in the rush period and 0 otherwise. It is entered linearly as were the previous variables. Thus, a separate factor is entered for the rush-hour period.

Age

Age was entered in 2 forms. Age was divided into 4 groups—less than 25, 26 to 45, 46 to 65, and older than 65—in an attempt to eliminate the linear effect of entering age directly into the model. This was done by using 3 dummy variables each taking either 1 or 0 value with a maximum of 1 variable having the value 1. It was hoped that this stratification would delineate different attitudes toward the separate modes at different age levels. Age was also tried as a linear variable.

Income

Income was treated in the same way as age. A linear correlation should not be expected between a variable such as this and choice, so that a set of 5 dummy variables was put forward. The groups were less than $5 thousand, $5 to $8 thousand, $8 to $12 thousand, $12 to $17 thousand, $17 to $25 thousand, and greater than $25 thousand per year. Some researchers have used income as a combining variable, especially with items of cost (for example, de Donnea, 25). Time limitations prevented that idea from being tested with respect to the multimode case.

Car Ownership

Car ownership was entered when an automobile was involved in the mode, i.e., for driving or riding. It was added linearly, but an argument could be made for making it a preemptive variable for the drive mode. That is, if the user does not own a car, he cannot possibly drive to the station.

Sex

The coefficient of the sex variable reflects the attitude of the female relative to the male toward the mode involved, for it is in the form of 0 or 1.

Stage in the Family Life Cycle

There was a high expectation for this set of variables. The set comprised 3 dummy variables that categorized the population sociologically according to 4 different classes. The first variable takes the value of 1 if the user is unmarried and living at home. The second has a value of 1 if he is unmarried and independent or if he is married with a spouse who does not compete for the use of the car. The third has a value of 1 if the user is married and has a spouse who goes to work separately. All others are together in the fourth group by reason that they will respond 0 to all of the above variables. The variables were entered to try to model the user characteristics and their relative traveling-mode priorities. It was felt that there might be some interrelation between this variable and the ownership variable.

Trip Purpose

The trip purpose variable could be used to describe the different economic demand generated by the separate trip purposes. That would involve stratifying and thus complicating the model, making it computationally impossible for the capacity of the program used unless separate models are built for each purpose. Thus, a constant demand function was assumed with respect to trip purpose, and the variable was not included.

THE RESULTS

Below is an outline of the most significant model obtained in runs of the maximum likelihood program to develop the logit model coefficients. Variables were eliminated if the t-value (in brackets below) obtained was not significant at the 0.90 level, and the program was rerun. Some of the variables (stage in the family life cycle and age) mentioned in the previous sector failed to be significant at all as a result.

For walk mode,

$$G_1(x) = 0 \text{ (base mode)}$$

For drive mode,

$$G_2(x) = \underset{(5.03,\,0.9995)}{-0.114\,\Delta C} - \underset{(5.14,\,0.9995)}{0.00421\,\Delta t} + \underset{(3.29,\,0.995)}{0.238\,\text{FRICT}} - \underset{(214,\,0.975)}{0.0123\,\text{WALK PLS}} + \underset{(3.32,\,0.995)}{1.49\,\text{CAR}}$$

$$+ \underset{(1.57,\,0.900)}{0.0108\,\text{AGE}} - \underset{(2.22,\,0.975)}{5.57\,\text{ID1}} - \underset{(3.28,\,0.99)}{7.42\,\text{ID2}} - \underset{(3.80,\,0.999)}{7.97\,\text{ID3}} - \underset{(3.53,\,0.995)}{7.09\,\text{ID4}}$$

For driven mode,

$$G_3(x) = 0.114\,\Delta C - 0.00421\,\Delta t - \underset{(3.45,\,0.995)}{2.26\,\text{RUSH}} + \underset{(1.97,\,0.950)}{1.169\,\text{ID4}}$$

For bus mode,

$$G_4(x) = -0.114\,\Delta C - 0.00421\,\Delta t - \underset{(1.85,\,0.950)}{1.34\,\text{RUSH}} + 2.025 \text{ (a constant)}$$

Likelihood ratio test, 120.06 with 15 deg of freedom; proportionate pseudo R-squared, 0.686

In the statements given above, ΔC is cost difference; Δt is time difference; ID1, ID2, ID3, ID4 are income dummy variables; FRICT is $\frac{1}{2}$ cost of parking + walking from parking at 6 cents/min; WALK PLS is the walking time in sec; CAR is car ownership; AGE is age; and RUSH is time of day.

The likelihood ratio test is highly significant even as high as 0.999, indicating that the model is valid. The proportionate pseudo R-squared statistic is only an approximation and not a true R-squared value. The latter value was used for comparison of models within the research program only and should be taken only as a rough guide and not as absolute.

A multiple F-value of 6.954 was obtained in a secondary evaluation of the model (on the larger data set). (A value of 3.38 is significant at the 0.999 level.)

From this set, the mean true proportions and those predicted by the model when the predicted probabilities were summed and when the exact count was made from the individual decisions were as follows:

Mode	Mean True	Summed Probabilities	Exact Count
Walk	0.135	0.112	0.078
Drive	0.293	0.217	0.208
Driven	0.469	0.398	0.453
Bus	0.103	0.273	0.191

This indicates that 17 percent were misplaced for the probabilities sum and 8 percent were misplaced for the absolute count. Unfortunately, for both cases, all of those misplaced were put on the buses. That indicates that $G_4(x)$ is overestimating and may be attributable in part to the constant being placed in that modal sector. Throughout the research, the placing of the constant had a definite effect on the model. Further research is required to investigate the sensitivity of the constant.

A secondary result of the model as constituted is the derivation of a value of time. In this case, it is a value of time saved. From the above coefficients of time difference and cost difference, we calculate the value $1.33/hour. Based on a 2,000-hour work year and the average yearly wage of the data set ($11,000), this value is 24.1 percent of the wage rate. That is only about half the value derived by Lisco (14) in his probit analysis. Stopher (22) found the value to vary from 0.33 to 0.14 depending on the salary. The above value falls in the middle of that range.

CONCLUSIONS

The most important conclusion that can be made is that it is possible to build a significant modal-split model by using this technique. Some further work should be done with respect to the constant.

The value of time for the short journey seems to be less than similar values derived by Lisco for longer trips in the same geographical area.

Many of the more definite conclusions concern the choice of variables for the models. The most disappointing variable was stage in the family life cycle, which failed to be significant in any of the sectors. Perhaps it would reflect a large importance when related to the longer overall trip.

Once again, time and cost proved to be significant as explanatory variables. That was expected. What is slightly surprising is that cost attained the same level of significance as time. It was expected from previous studies that time would be the more important variable. The difference in significance between the two, however, was only minimal. It can, therefore, be concluded that time and cost are both highly important in consideration of this modal choice.

One convenient advantage of this method is the unavailable mode situation. If, for example, the bus were not available to a user, it could be assigned an arbitrarily large time difference so that its probability were reduced to a very small value approaching zero.

It is not possible to claim that this is an operational model. The variable set should be refined, the problems with a constant should be studied, and a definite statistical advantage should be established over other modeling techniques. If the evening peak is used, there may be a whole different set of variables, for the results of the morning peak put definite limits on the evening return trip. Given that the problems can be resolved, the next step is an extension to include the other 3 steps in the urban transportation planning package.

ACKNOWLEDGMENT

This paper is the result of research conducted in the Department of Civil Engineering and Engineering Mechanics at McMaster University, Hamilton, Ontario, under the supervision of Peter Stopher. The data base used in the research was collected in suburban north Chicago by the Chicago Area Transportation Study.

REFERENCES

1. Adams, W. T. Factors Influencing Transit and Automobile Use in Urban Areas. Public Roads, Vol. 30, 1959.
2. Betak, J., and Betak, C. Urban Modal Choice: A Critical Review of the Role of Behavioural Decision Variables. Chicago, 1969.
3. Biciunnas, A. E. Modal Split of First Work Trip. CATS Res. News, Chicago Area Transportation Study, Vol. 6, No. 1, 1964.
4. Carroll, J. D. Relation Between Land Use and Traffic Generation. BPR Res. Digest, No. 1, Feb. 1954.
5. Davis, H. E. Urban Transportation Planning: Introduction to Methodology. In Defining Transportation Requirements, ASME, 1969.
6. de Donneau, F. Probit and Logit Modal Split Models for Rotterdam. Netherlands Economic Institute, Rotterdam, Dec. 1970.
7. Deen, T., Mertz, W., and Irwin, N. Application of a Modal Split Model to Travel Estimates for the Washington Area. Highway Research Record 38, 1963, pp. 97-123.
8. Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. Ann. Eugen. Lond., Vol. 7, 1936.
9. Hadden, J. K. The Use of Public Transportation in Milwaukee. Traffic Quarterly, Vol. 18, 1962.
10. Hamburg, J. R., and Guinn, C. R. A Modal Split Model: Description of Basic Concepts. New York State Dept. of Public Works, Albany, 1966.
11. Hill, D. M., and Von Cube, H. G. Development of a Model for Forecasting Travel Mode Choice in Urban Areas. Highway Research Record 38, 1963, pp. 78-96.
12. Howe, J. J. Modal Split on CBD Trips. CATS Res. News, Chicago Area Transportation Study, Vol. 2, No. 12, 1958.
13. Lave, C. A. Modal Choice in Urban Transportation: A Behavioral Approach. Stanford University, PhD dissertation, 1968.
14. Lisco, T. E. Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, PhD dissertation, 1967.
15. Martin, B. V., Memmott, F. W., and Bone, A. J. Principles and Techniques of Predicting Future Demand for Urban Area Transportation. M.I.T., Cambridge, Res. Rept. 38, June 1961.
16. McGillivray, R. G. Binary Choice of Transport Mode in the San Francisco Bay Area. Univ. of California, Berkeley, PhD thesis, 1969
17. Quarmby, D. A. Choice of Travel Mode for the Journey to Work: Some Findings. Jour. of Transp. Econ. and Policy, Vol. 1, No. 3, 1967, pp. 273-314.
18. Quinby, H. D. Traffic Distribution Forecasts: Highway and Transit. Traffic Eng., Vol. 31, No. 5, 1961.
19. Rassam, P. R., Ellis, R. H., and Bennett, J. C. The N-Dimensional Logit Model: Development and Application. Peat, Marwick, Mitchell and Co., Washington, D.C., unpublished, 1970.
20. Sharkey, R. H. A Comparison of Modal Stratification of Trips by Distance From the CBD. CATS Res. News, Chicago Area Transportation Study, Vol. 2, No. 5, 1958.

21. Sharkey, R. H. The Effect of Land Use and Other Variables on Mass Transit Usage in the CATS Area. CATS Res. News, Chicago Area Transportation Study, Vol. 3, No. 1, 1959.
22. Stopher, P. R. Report on the Journey to Work Survey. Res. memo, London, Jan. 1968.
23. Stopher, P. R. A Probability Model of Travel Mode Choice for the Work Journey. Highway Research Record 283, 1969, pp. 57-65.
24. Thiel, H. A Multinomial Extension of the Linear Logit Model. Internat. Econ. Rev., 1969.
25. Warner, S. L. Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice. Northwestern Univ. Press, Evanston, Ill., 1962.
26. Wynn, F. H. Land Use Relationships With Travel Patterns. BPR Res. Digest, Nov. 1955.

SUBJECTIVE TIME SAVINGS IN INTERURBAN TRAVEL: AN EMPIRICAL STUDY

Shalom Reichman, Hebrew University of Jerusalem

Travel time savings are usually estimated on the basis of objective measures of time. Such objective measures are assumed to correspond to the mean of a distribution of subjective time savings as reported by travelers. In a multimodal, single-route, interurban passenger survey carried out in Israel, that was not found to be the case. Each passenger was requested to report separately time differences by mode and time savings. When the replies were compared, it appeared that 21 percent of air passengers stated that their time savings amounted to quantities nearly twice as much as the mean of the difference in reported travel times. When asked how much time they would have saved traveling by air, 16 percent of all bus passengers indicated the same discrepancy between differences in time spent and time saved. A group of large time savers, as compared with normal time savers, was identified with a distinctive profile of trip attributes but not of socioeconomic attributes. Large time savers are those passengers who reported twice as much time savings as the differences in time spent by mode. Possible reasons for the bimodal distribution of time savings have been sought in the perception of the choice situation generally facing the traveler, or, alternatively, in the special conditions of trips across a desert. It is recommended to clarify the generality of the results by means of additional surveys of similar design.

THE COMMON observable fact that time has opportunity costs is probably best seen in the purchase of fast but costlier transportation services. Not surprisingly, mode choice studies have been used as a prime tool to determine the value of travel time as perceived by the individual decision-maker.

The basic assumption is that each transportation technology has 2 prime choice attributes: speed and cost. If for every journey having a given origin and destination at least 2 different routes or modes can be chosen, then, other things being equal, the selection will be made on the basis of the value of time of the traveler. People with a low value of time will choose the slower route or mode, and people with the higher value of time will choose the faster and more expensive route or mode.

Additional assumptions relate to the strict applicability of market place conditions to the travel mode choice situation. Rationality in the choice-making, full-information, and freedom of choice are usually included here. A more intractable problem concerns the fact that transportation services are actually joint products, combining capacity and quality of service. So far, no way has been found to treat the comfort side of the product adequately.

The present study is concerned with the empirical investigation of the assumption that interview respondents are able to perceive accurately the objective travel time that they face in a choice situation. Stated alternatively, the same assumption claims that there is a consistent relationship between objective time data as derived from engineering assignments and the perceived estimates of time consumed in travel. The

Publication of this paper sponsored by Committee on Social, Economic and Environmental Factors of Transportation.

simplest situation would be that in which random errors in estimation are distributed normally around an average close to the objective time. In the case of the Skokie study (1, 2), it has been found that such simple situations do indeed occur, when travelers were asked on time spent on a commuter trip, though without being asked as to their estimate of time differences or time savings by mode.

It is suggested that estimations of travel time by travelers differ from objectively measured times in at least one aspect. Whereas measured times can be obtained repeatedly within a narrow confidence limit, irrespective of the distance traveled, travel time estimations by travelers might be affected by trip distances. Time estimations of interurban trips, unlike commuting, may be systematically biased for a variety of reasons: the perception of distant locations in the real world, preference scales of the various transportation modes, and time savings of greater magnitude measured by hours rather than fractions of hours. All of these could cause significant and consistent discrepancies between measured and estimated travel times.

In this paper, "subjective" time refers both to travel time differences by mode and to travel time savings, provided that both are based on estimations by travelers. "Objective" time, on the other hand, is determined by the unbiased measurement of the time consumed by mode on the basis of the performance of each mode. The point that will be made is that subjective time differences may differ substantially from subjective time savings. Consequently, a significant error may be introduced if estimations of travel time savings are based on objective measurements of differences in time consumed by mode.

More specifically, if subjective travel time savings estimates were to differ substantially from objectively measured time savings, the current state of the art in determining the value of travel time would be affected (3). Moreover, recent findings (4, 5) seem to indicate that the value of travel time is a function of the amount of time saved. Biased travel time estimations would then result in even greater errors in determining the value of time saved.

FINDINGS OF PREVIOUS STUDIES

The first indication of significant differences between objectively measured and subjectively reported time savings was found in a week-long air passenger survey carried out in Israel in November 1967 (6). All air travelers between Tel Aviv and Elat were asked to report on a questionnaire how much time, in their opinion, they saved by traveling by air instead of by ground transportation. For an adjusted subsample of travelers with origins and destinations within 5 miles of the airports, objective time savings were determined to average between 5 and 6 hours.

Rather than the expected normal frequency distribution or replies, averaging around 5 to 6 hours, the results showed a clear bimodal frequency distribution. About 50 percent of the travelers replied that they saved amounts similar to those that may be surmised from a probability distribution curve around the objective time difference between the modes. However, the other half of the subsample population replied that it had saved twice as much time, namely, a day or more. It should be noted that no 1-day round trips were included in the subsample.

In the subsequent analysis of the results, a number of questions could not be satisfactorily resolved because of the lack of sufficient data. The interpretation of possible reasons for the bimodal distribution was inconclusive, particularly because a number of alternative explanations could be made. To begin with, the information basis of the interviewed passengers was unknown. There was no indication whether the passengers had a correct knowledge of the objective times involved; hence, there is the possible existence of a perception bias (7). Similarly, even if it were assumed that the information basis was adequate, there still remained a reasonable explanation, in terms of an antimodal bias, for the excessive amount of time saved reported. In other words, because air travelers completed all questionnaires, they may have wished to exaggerate the advantages of the chosen over the rejected mode.

A repetition of the inquiry in the field was, clearly, both necessitated and justified by the partial results of the 1967 survey. This time, however, special attention was to be given to the design of the questionnaire and of the sample. In this way, it was hoped to isolate the effects of the perception and antimodal biases from other possible explanations for discrepancies between measured and estimated travel time savings.

THE SURVEY OF NOVEMBER 1970

In November 1970, the Israel Ministry of Transport carried out a field survey to determine the price of time as perceived by interurban passengers (8). As before, the choice situation consisted of mode, rather than route selection: All persons traveling between Elat and the rest of the country were required to complete a questionnaire. The alignment of bus routes was kept as close as possible to reality; thus, the number of buses leaving for Tel Aviv is larger than the number going in the other direction (Fig. 1). The survey lasted 7 days, but the seventh day had subsequently to be discounted because flash floods disrupted the single overland transportation link. The total number of questionnaires returned in 6 days was about 7,000; about 6,000 were suitable for analysis (9). The rate of response reached 90 percent.

Two main modifications were introduced in the 1970 survey. First, 3 modes were investigated instead of 1: all regular air services, all regular bus services, and all occupants of private vehicles. Modal split of the sample population is given in Table 1.

Second, the questionnaire included more than a single question relating to time estimation. One question remained roughly in its former wording. Only this time, bus passengers as well as air passengers were asked how much time they thought they would have saved by flying instead of going by bus. The second set of questions intended to reveal the information basis of the various time attributes by mode.

Hence, the following question was asked, How long do you think it would take to make this trip by bus, by car, by air? Passengers were asked to answer to all listed modes. It was assumed that, by the subtraction of time consumed by air from time consumed by bus or car, the information basis of the decision-maker as to mode attributes and spatial perception would emerge independently from subjective estimations of time savings reported in the first question.

DISCUSSION OF FINDINGS

The principal hypotheses with regard to the expected distribution of replies may be stated as follows:

1. The distribution of replies related to time saving estimation will be bimodal, irrespective of the means of transportation used—the first mode will average around 5 hours, and the second will be twice that magnitude or more; and
2. The distribution of replies to the question on travel time estimation by mode (or information variable) is likely to be normal for each transportation mode, with a possible shift of the value of the mean according to the transportation mode used.

Preliminary results (Tables 2 and 3) indicate that both hypotheses may be accepted. Of particular interest are the frequencies of replies in the category of 12 and more hours (Table 2). Disregarding those passengers who did not reply, 21 percent of total air passengers and 16 percent of total bus passengers stated that they saved (or would have saved) 12 or more hours. In the subsample of Elat residents, that proportion is even higher, 28 and 23 percent respectively.

The results (Table 3 and Fig. 2) reveal the existence of a normal distribution, with a shift of the mode and median according to the transportation mode used. Thus, the mode for air passengers is 5 hours of time differences between air and bus, and the median is slightly more. For bus passengers, the mode is also 5 hours, and the median is slightly less. Car passengers, as expected, estimated time differences to be far less; the mode is only 4 hours, and the median is even less. An interesting corroboration of these antimodal biases is found in the distribution replies by the Elat residents, who are assumed to have a greater knowledge and experience regarding the times involved. In all cases, the differences as given by Elatis accentuate the mode and median difference but do not affect the basically normal distribution.

Figure 1. Transportation services and routes to Elat.

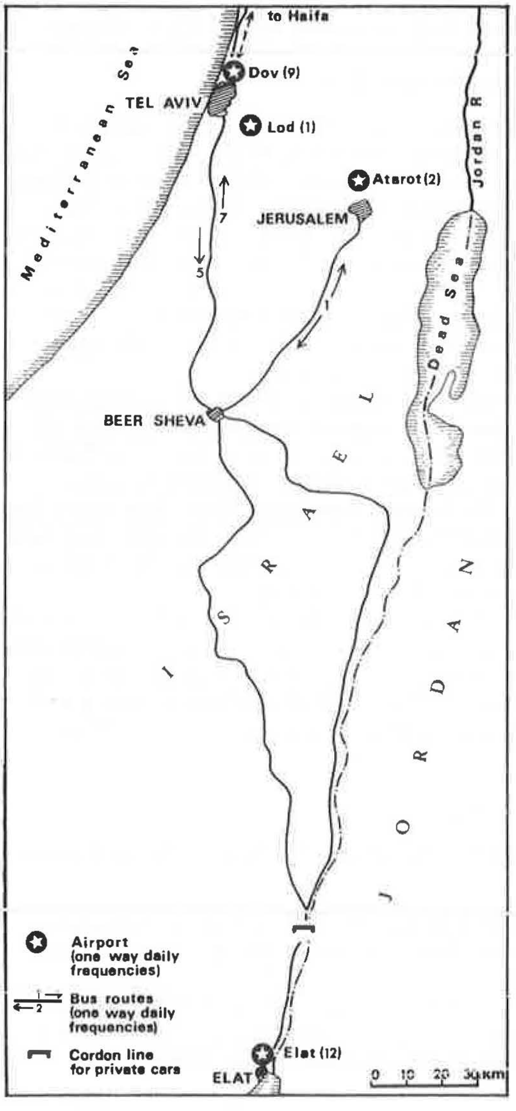
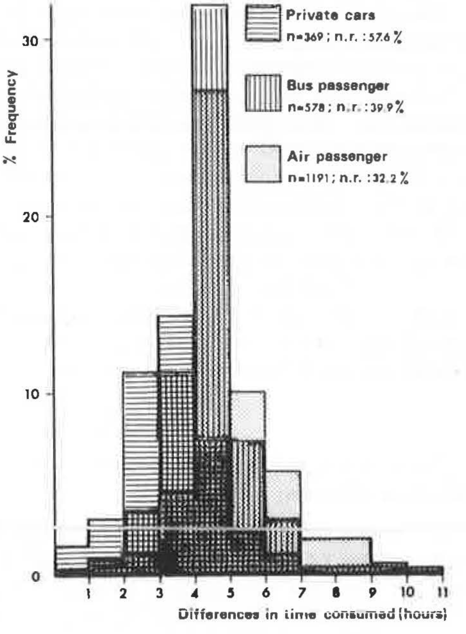


Table 1. Modal split of Elat survey.

	Total Replies		Usable Replies	
Mode	Number	Percent	Number	Percent
Air	4,036	57	3,811	64
Bus	2,141	31	1,451	26
Private vehicles	962	12	644	10
Total	7,139	100	5,906	100

Note: Does not include Saturday, October 31, because no public transportation operates on Saturday. Also, replies given on November 6 were not included, because overland routes were flooded and traffic was interrupted.

Figure 2. Passengers' estimation of differences in time by transportation mode.



The data were subsequently checked to ascertain that respondents replied to both time-related questions—the first about estimates of time saved and the second about estimates of time consumed by each mode. On the basis of this analysis, 2 populations were separated: group A, a subpopulation that reported time savings of similar magnitudes to the differences in time consumed; and group B, a subpopulation that reported time savings far in excess of their reported time differences by mode. More specifically, group A included all respondents with savings of 12 hours or fewer, and group B included all respondents who estimated their time savings as more than 12 hours, a day, and more than a day. Group A included 3,255 respondents, and Group B included 821 respondents. Travelers who did not reply to both questions were excluded at this stage.

It is proposed that the existence of the 2 populations can be related to the nature of transportation demand, sometimes defined as derived demand (10). The purpose of a trip, which is one of the main determinations of the decision as to whether to make a displacement, is usually to perform another activity at the trip end. Thus, it is possible that decision-makers might assess the savings of time of a trip not just in terms of the attributes of a given transportation technology, such as speed, but also in terms of the amount of time left at either trip end for the completion of the activity that generated the trip. For example, in the case of the Elat- Tel Aviv route, a 5-hour trip by surface transportation will in some cases mean the disruption of a full working day. The distinction between the 2 groups of travelers might therefore lie in the perception of the choice situation when travel time can be saved.

Most people might perceive the choice in the strict sense, that is, between 2 alternative technologies. For the other group, the choice situation is broader because other activities, besides traveling, are being considered. Whenever technological time savings become large, such as 4 to 6 hours as in this case, it can be assumed that the difference between the 2 perceptions of the choice situation would be substantial.

Another reason for the large difference in time saving could be the result of spatial perception biases rather than perception of choice-making situations. The trip to Elat involves the crossing of a large desert area, so that travelers might conceivably feel that the advantage of flying over the desert rather than driving across it is equivalent to the saving of much more time. The general case would be that origins and destinations of trips might affect the estimation of time savings.

On the basis of these propositions, each group was characterized by a profile consisting of 3 elements: socioeconomic properties, trip properties, and spatial properties. It is suggested that large time savers, namely group B, have distinctive profiles of trip and spatial characteristics, though not necessarily in their socioeconomic properties. The 2 profiles are given, by means of selected representative values, in Table 4.

It appears that the 2 groups have similar socioeconomic characteristics but show large differences in their trip characteristics. As for spatial properties, both properties seem to differ although that of the general origin and destination distribution of trips is slightly less significant. A typical profile or a large time saver is likely to be a resident of Elat who is traveling by air to a large city for work purposes and who, preferably though not necessarily, will return on the same day.

CONCLUSIONS

On the basis of the empirical results analyzed so far, it has been shown that differences between objective and subjective amounts of time saved do exist. Those differences are not just the expression of information or antimodal perception biases, although those have been found to occur. A group of passengers has been identified that consistently estimated time savings at least twice as large as the usual objective differences in travel times.

It may be tempting to discuss the implications of these results for modal-split modeling, disaggregate behavior models, or the evaluation of travel time savings. However, it is felt that such an interpretation of the results is premature.

Two problems in particular await further clarification by additional surveys of similar design in other parts of the world. First, it has to be established whether the

Table 2. Time savings by transportation mode and residence.

Hours	Air Passengers				Bus Passengers			
	All		Elat Residents		All		Elat Residents	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
0 to 2	112	3.0	36	3.6	57	3.9	12	3.7
2 to 4	327	8.8	77	7.6	146	10.1	37	11.4
4 to 6	829	22.1	192	18.9	380	26.2	72	22.2
6 to 9	346	9.3	68	6.7	70	4.8	12	3.7
9 to 12	206	5.6	43	4.2	57	3.9	7	2.2
12 to 18	28	0.8	5	0.5	6	0.4	0	0.0
18 to 24	18	0.5	4	0.4	6	0.4	1	0.3
24 to 48	184	4.9	73	7.2	71	4.9	18	5.6
None specified								
Less than a day	534	14.3	146	14.4	222	15.3	34	10.5
More than a day	405	10.9	139	13.7	97	6.7	34	10.5
No reply	738	19.8	231	22.8	339	23.4	97	29.9
Total	3,727	100.0	1,014	100.0	1,451	100.0	324	100.0

Note: Time saving estimation was derived directly from the question, How much time do you think you saved (for bus, you would have saved) by flying instead of going by bus? For reasons of comparability with previous surveys, passengers with origins or destination in Sinai have been excluded.

Table 3. Differences in time by transportation mode and residence.

Difference in Time by Air and Bus (hours)	Air Passengers				Bus Passengers				Private Car			
	All		Elat Residents		All		Elat Residents		All		Elat Residents	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
0	9	0.2	6	0.6	0	0.0	0	0.0	2	0.3	0	0.0
1	17	0.5	1	0.1	3	0.2	1	0.3	8	1.2	2	1.6
2	42	1.1	21	2.1	12	0.8	3	0.9	20	3.1	5	4.0
3	168	4.6	53	5.2	55	3.8	17	5.2	74	11.5	16	12.7
4	471	12.7	158	15.6	165	11.4	42	12.9	95	14.8	14	11.1
5	1,019	27.4	306	30.2	465	32.1	91	28.2	49	7.6	5	4.0
6	383	20.4	14	11.2	111	7.6	26	8.0	17	2.6	2	1.6
7	219	5.9	33	3.6	45	3.1	8	2.5	8	1.2	0	0.0
8	76	2.0	9	0.9	6	0.4	1	0.3	2	0.3	0	0.0
9	73	2.0	5	0.5	3	0.2	0	0.0	0	0.0	0	0.0
10	27	0.7	3	0.3	5	0.3	2	0.6	0	0.0	0	0.0
11	15	0.4	2	0.2	3	0.2	1	0.3	0	0.0	0	0.0
12	5	0.1	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
13	3	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
14	2	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
15	3	0.0	1	0.1	0	0.0	0	0.0	0	0.0	0	0.0
16	2	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
17	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
18	2	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
No reply	1,191	32.0	298	29.4	578	39.9	132	40.8	369	57.4	82	65.0
Total	3,727	100.0	1,014	100.0	1,451	100.0	324	100.0	644	100.0	126	100.0

Note: Computed from the question, How long do you think it would take to make this trip by bus, by air, by car? and subtracting the time of the fastest mode from the slowest.

Table 4. Profiles of time-saving groups by main properties.

Item	Group A	Group B
Number of cases	3,255	821
Socioeconomic properties		
Age 22 to 40, percent	59.1	60.4[a]
Monthly income I£ 601 to 2,000, percent	47.9	52.6[a]
Trip properties		
Bus travelers, percent	37.8	22.2[b]
Work-recreation ratio	0.56	0.89[b]
Return same day, percent	12.1	24.2[b]
Spatial properties		
Elat residents, percent	27.4	38.2[b]
Origins and destinations in Tel Aviv, Jerusalem, Haifa, percent	84.0	81.1[c]

[a] Difference between groups not significant.
[b] Difference significant at 99.91 percent.
[c] Difference significant at 99 percent.

group of large time savers found in the Elat case represent a general case of the perception of the choice situation in a broader sense or merely represent a unique example of a particular spatial bias affecting desert towns. The other problem relates to the stability of the perception over time and space. Here again, highly divergent outcomes may be hypothesized and, therefore, require further investigation.

REFERENCES

1. Harrison, A. S., and Quarmby, D. A. The Value of Time in Transport Planning: A Review. Economic Research Center, European Conf. of Ministers of Transport, Paris, 1969.
2. Lisco, T. E. The Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, PhD thesis, 1967.
3. Gronau, R. The Effect of Travelling Time on the Demand for Passenger Airline Transportation. Univ. of Columbia, PhD thesis, 1967.
4. Thomas, T. C., and Thompson, G. I. The Value of Time for Commuting Motorists as a Function of Their Income Level and Amount of Time Saved. Highway Research Record 314, 1970, pp. 1-19.
5. Thomas, T. C., and Thompson, G. I. The Value of Time Saved by Trip Purposes. Highway Research Record 369, 1971, pp. 104-117.
6. Reichman, S. Passengers on Arkia Services in November 1967. Jerusalem, Aug. 1968.
7. Watson, P. Problems Associated With Time and Cost Data Used in Travel Choice Modeling and Valuation of Time. Highway Research Record 369, 1971, pp. 148-158.
8. The Value of Time Saved in Domestic Aviation. Division of Planning and Economics, Ministry of Transport, Jerusalem, Int. Rept., July 1971.
9. Reichman, S. Spatial Characteristics of Short-Haul Air Traffic: The Case of the Tel-Aviv Route. Paper presented at meeting of Commission of Transportation Geography, International Geographical Union, Toronto, July 1972.
10. Oi, W. Y., and Shuldiner, P. W. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, Ill., 1962.

PREDICTIONS OF INTERCITY MODAL CHOICE FROM DISAGGREGATE, BEHAVIORAL, STOCHASTIC MODELS

Peter L. Watson, Transportation Center, Northwestern University

This paper reports the results of some empirical prediction tests on disaggregate, behavioral, stochastic models of modal choice. It is discovered that the expected-number calculation (summation of probabilities) yields modal-choice predictions with a high degree of accuracy, although a caveat is in order regarding the transferability of such models: They should be transferred with care and only to similar situations. The stability of the models is evidenced by the fact that they can be broken down by trip purpose and reaggregated without loss of predictive power. Finally, the method of obtaining predictions by classifying probabilities is considered and rejected.

•IN RECENT years, work in the area of modal choice has tended to concentrate on the development of disaggregate, behavioral, and stochastic models. (Modal choice should be interpreted as including route choice.) Those developments have led to a model type that has a number of advantages over the traditional zonal-based models. First, the building of models at the disaggregate level with the individual as the basic decision-making unit means that problems of zonal homogeneity are avoided and that better use is made of the data collected. Second, because the models are based on the observable behavior of the individual traveler, they are more realistic. Third, the introduction of the stochastic element means that modelers are now concerned with the probability that a given mode will be chosen rather than with the modal split of a set of travelers from a zone. [A more detailed and comprehensive statement of the advantages of disaggregate, behavioral, and stochastic models is given by Stopher and Lisco (1).] The result of these modifications is that a generation of models has evolved that perform most successfully in terms of their ability to describe and explain modal choices. The models both reflect acceptable hypotheses about traveler behavior and achieve acceptable levels of significance in a statistical sense.

However, if the new generation of models is to be a useful addition to the planner's set of tools, it must be demonstrated that the new models have acceptable prediction properties. A comparison of the prediction capabilities of both aggregate and disaggregate models would be desirable, but documented evidence on the aggregate models is not readily available. Moreover, the aggregate nature of the data used in zonal-based models means that they cannot be used to produce disaggregate predictions in the same situation. Until the 2 types of model can be tested on the same data set, one-sided statements of prediction capability must suffice.

It is the aim of this paper to investigate the capabilities of disaggregate, behavioral, and stochastic models of modal choice. An exposition of the prediction methodology is followed by a series of empirical tests that determine predictive performance under different conditions.

The tests were carried out by using data from the Edinburgh-Glasgow area modal-split study, which investigated modal choices in the Forth-Clyde corridor of the central lowlands of Scotland. Details of the study background, data collection, and model

Publication of this paper sponsored by Committee on Social, Economic and Environmental Factors of Transportation.

development are given in another report (2). The travel data represent medium-range intercity trips. The choice modeled is between car and train; each of those modes carried approximately 45 percent of total traffic. The main aim of the study was to examine the transferability of disaggregate models from intracity commuting situations to other trip lengths and purposes; thus, models were developed for 3 trip purposes—journey to and from work (JTW), business (BUS), and social-recreational (SOCREC)—and for total travel. The variables that resulted in the best explanatory models for each trip purpose are given in Table 1. The abbreviations in Table 1 are defined as follows:

TD Rel = time difference relative to total journey time,
CD Rel = cost difference relative to total journey time,
WW Tim = difference in walking-waiting time,
WW Rel = difference in walking-waiting time relative to total journey time,
TJT Ca = total journey time by car,
SUBCOS = cost of access-egress to and from station,
Ju Tra = number of walk, wait, and travel segments in train journey, and
Ju Diff = difference in number of segments by each alternative.

The "relative" versions of time and cost differences and the relative walk-wait time variable reflect the fact that a given time (or cost) saving becomes less important as the length (or cost) of the journey increases. The "journey unit" variables represent a variable developed to reflect the inconvenience of each trip by allocating 1 trip unit to each segment (walking, waiting, and riding) of the trip.

The models were estimated by using logit analysis (3) so that

$$P_t = \frac{e^{G(x)}}{1 + e^{G(x)}} \tag{1}$$

where P_t is the probability of choosing the train, and G(x) is a linear combination of the explanatory variables.

METHODOLOGY

Two main methods of obtaining predictions from disaggregate models of modal choice have been advocated. The first proceeds by a summation of predicted choice probabilities, and the second is a classificatory approach. The following discussion of the methods of obtaining predictions is set in terms of a car-train choice, where the dependent variable is the probability that the train will be chosen. Having calibrated the model, one can calculate the probability of choosing the train for each individual traveler, and those probability estimates are used to derive the prediction of the number who will choose the train. The predictions are derived on the assumption that the important result from a prediction point of view is the expected number of travelers who will choose a given mode. Thus, the models are calibrated by using the individual traveler as the basic unit of analysis, and the results are then aggregated to produce results that are meaningful from a planning standpoint. (The appropriateness of this approach will be taken up again in the discussion of the criteria that will be used to evaluate the prediction results.)

Table 1. Model content.

Purpose	Variables
JTW	TD Rel, CD Rel, WW Rel, Ju Dif
BUS	TD Rel, CD Rel, WW Tim, Ju Dif
SOCREC	TJT Ca, SUBCOS, WW Tim, Ju TRA
Total	TD Rel, CD Rel, WW Tim, Ju Dif

Expected-Number Approach

This method derives from the fact that the expected value of a random variable is the sum of the products of the potential outcomes and the probabilities of their occurrence. Binary outcomes, coded as 0 or 1, are the special case of a Bernoulli random variable such that

$$E(p) = 1\,(P_T) + 0\,(1 - P_T) \tag{2}$$

$E(p)$ = expected value of the probability, P_T = probability of choosing the train, and $1 - P_T$ = probability of not choosing the train (i.e., of choosing the car). Thus, the expected value is equal to the probability of choosing the train. To obtain the expected number of travelers who will choose the train (i.e., the expected number of positive outcomes in the sample) requires that the expected values be summed.

$$E(\#) = \sum_{i} E(P_{Ti}) = \sum_{i} P_{Ti} \tag{3}$$

$E(\#)$ = expected number of positive outcomes (train choices), and P_{Ti} = probability that the i th traveler will choose the train.

An example may clarify this point. Suppose that 10 travelers had the following observed probabilities and choices, where 1 = choice of train:

Choice	Predicted Probability
1	0.9
1	0.7
0	0.2
0	0.5
1	0.8
0	0.4
1	0.8
1	0.6
0	0.3
1	0.7

In this example, 6 travelers chose the train, i.e., there were 6 travelers with a probability of choosing the train equal to one; the sum of those probabilities is equal to the number of travelers choosing the train. The prediction of the number of train travelers (the expected number) is obtained by summing the predicted probabilities. That sum is equal to 5.9; therefore, it is predicted that 6 travelers will choose the train.

Classification Approach

This method is based on the premise that the predicted probabilities can be used to assign travelers to the alternative modes. Misclassifications are minimized by allocating travelers about a probability of 0.5. This means that, if the traveler's probability of choosing a train is greater than 0.5, he is assigned to the set of train travelers; if the probability is less than 0.5, he is assigned to the set of car travelers. The numbers assigned to each set then constitute the prediction of the number who will choose each alternative.

Discussion of Methods

To examine the relative performance of these 2 methods requires the establishment of criteria that will be used to judge them. Although the models that yield the predictions of probabilities are disaggregate models, in the sense that the unit of analysis is the individual traveler, testing their predictive abilities at a disaggregate level is inappropriate for a number of reasons. The true probability that a given traveler will choose the train is unknown. It is, therefore, not possible to test the ability of the models to predict (accurately) the probability of a given choice. The known information is the choice that the traveler was observed to make, and it is possible to compare the predicted probability with the observed choice. That is the basis of the classification criterion that has been applied to predictions by classification. However, a consideration of the meaning of the probability of choice reveals that the use of classificatory ability as a test of a model that predicts probabilities is a meaningless test. Consider, for example, a traveler for whom the probability of choosing the train is

0.6. That probability can be interpreted as implying that, for a large number of observations, the traveler would choose the train 60 percent of the time. The classification approach would, however, assign him to the train group and, if he were observed to take the car, would treat it as an erroneous classification.

Thus, neither the expected-number approach nor the classification approach should be judged on its ability to predict either the probabilities or the actions of individual travelers. That is not a serious problem, however, because the use of disaggregate models in a planning context requires that aggregate, not disaggregate, predictions be derived. This leads to an evaluation of both the prediction approaches and the predictive capabilities of the disaggregate models themselves in terms of their abilities to produce accurate aggregate predictions of the number of travelers who will choose a given mode.

PREDICTION TESTS

Although the procedures that are used to derive predictions are very simple, a caveat is in order before the prediction tests are undertaken. The most appropriate test of a predictive model involves 2 data collection efforts: one to calibrate the model and the other, preferably after a system change, to test the predictions. Such an effort is seldom possible. Thus, it is necessary to employ a second-best approach in which both the model and the predictions must be tested on the same data set. The more ideal situation of 2 random samples is approximated from the population by the following procedure. The data set is randomly divided into halves: One half is used to calibrate the models, and the other is used for the prediction tests. The tests described in this paper employ a double-edged version of this procedure in which coefficients are estimated for both halves of the data set, after which the coefficients and data sets are interchanged to obtain 2 sets of predictions, which are summed to provide the final prediction. It is acknowledged that this testing procedure is not ideal, but, in the absence of more extensive data collection facilities, it is the best available.

After the models given in Table 1 are calibrated, the expected-number procedure was used first to obtain predictions of the expected number of train travelers. The results are given in Table 2. It is clear that the differences between the number of travelers who are observed to choose the train and the number who are predicted to choose the train are small. Two tests are proposed that will indicate the magnitude of the prediction errors. The first considers the difference between the predicted and observed numbers of train travelers and expresses the error in prediction as this difference relative to the observed number. In other words, ϵ_1 indicates how close the model comes to providing a perfect prediction of the number of train travelers. The second takes a rather broader view of the prediction process and examines it from the point of view of modal split. Given the nature of the predictions, i.e., that the probability of choosing the train, P_T, is equal to one minus the probability of choosing the car, $1 - P_T$, the models also provide a prediction of the number of travelers who will choose the car, $\sum_i (1 - P_{Ti})$. It is obvious that the difference between predicted and

Table 2. Predictions of number choosing train.

Approach	Purpose	Set 1 Predicted	Set 1 Observed	Set 2 Predicted	Set 2 Observed	Sets 1 and 2 Predicted	Sets 1 and 2 Observed	Number
Expected number	JTW	125	117	123	135	248	252	360
	BUS	245	249	241	235	486	484	878
	SOCREC	241	237	208	211	449	448	1,202
	Total	582	589	601	595	1,183	1,184	2,440
Classification	JTW	138	117	143	135	281	248	360
	BUS	263	249	260	235	523	484	878
	SOCREC	148	237	182	211	330	448	1,202
	Total	574	589	609	595	1,183	1,184	2,440

observed train travelers will be equal (although a different sign) to the difference between predicted and observed car travelers. That being the case, the number of erroneous predictions can be regarded as an indicator of the failure of the model to predict correctly the modal split and may be expressed relative to the sample size to provide an alternative measure of prediction performance ϵ_2. Thus,

$$\epsilon_1 = \frac{\text{OBS - PRED}}{\text{OBS}} \times 100$$

and

$$\epsilon_2 = \frac{\text{OBS - PRED}}{\text{N}} \times 100$$

Applying those error measures to the predictions given in Table 2 yields the following percentages:

Measure	JTW	BUS	SOCREC	Total
ϵ_1	1.59	0.41	0.22	0.08
ϵ_2	1.11	0.23	0.08	0.04

Although it may be argued that ϵ_2 overstates the performance of those models, it is clear that they predict extremely well.

Next, the classification approach was used to obtain predictions of the number of train travelers. These predictions are essentially different from those presented in the preceding section because they derive not from the summation of probabilities but from the use of the estimated probabilities to assign travelers to modes. Thus, the basis of this method is a classification procedure that assigns travelers with an estimated probability of more than 0.5 to the train and those with an estimated probability of less than 0.5 to the car. It should be noted at this point that the estimated probabilities that are used were obtained from the random-division estimations. The resulting predictions are also given in Table 2.

The error percentages associated with the predictions by classification are as follows:

Measure	JTW	BUS	SOCREC	Total
ϵ_1	13.31	8.06	26.34	0.08
ϵ_2	9.17	4.44	9.82	0.04

In evaluations of the predictive performance of these models, both absolute and comparative with regard to the alternative methods of obtaining the predictions, it is important to note that the tests carried out reflect the ability of the models not simply to replicate the data on which they were calibrated (the nature of the calibration technique means that such a procedure yields perfect replications) but to predict travelers from a second random sample that is taken from the same population as the sample used to calibrate the model.

In the light of this consideration, the predictions obtained from the expected-number approach are remarkably accurate. If only the ϵ_1's are considered, the highest error among the 4 models is 1.59 percent. The classification approach does not perform so well, and the errors associated with the predictions derived by classification are several orders of magnitude greater than those by the expected-number procedure.

Several factors may contribute to the relative failure of the classification approach. It was argued above that classification implies the ability to make precise inferences about an individual's action from his predicted probability. However, probability statements are only meaningful from a prediction point of view when aggregated either for the population of travelers or for multiple trips by the individual traveler. It seems very likely that the failure of the classification method to result in accurate predictions is a result of the inappropriate attempt to match predicted probabilities with specific actions.

PREDICTIONS WITH NONRANDOM DIVISIONS

It has been argued that the random-division procedure is improper when large data sets are used because the new distributions should differ little from the parent distribution and, therefore, the process of randomly dividing the data should result in little change. It is against that background that the non-random-division tests are advocated. The argument that nonrandom divisions should be used to counter the above problems is not very attractive because such divisions are usually based on income (or other socioeconomic characteristics). If one believes that those factors affect modal choice, it is unreasonable to expect that a model built for one income group would predict well for travelers in a different income group. In short, it is argued that a model should be able to predict well in the case of a further sample drawn from the original population, (i.e., a sample with similar distributional characteristics to the one used to calibrate the model); it is not reasonable to expect a model to predict well when presented with a new set of data whose distributional characteristics are different from those of the calibration data, whether the differences are the result of non-random-resampling procedures or even the transfer of the model to a new, but different, situation.

COMPOSITION OF THE PREDICTIONS

Having undertaken the above detour to deal with the question of nonrandom divisions, it is now appropriate to return to the main stream of this paper. The prediction results presented above were obtained for 3 trip purposes—journey to and from work, business, and social-recreational—and for total travel. It should be emphasized at this point that a different model was constructed for each trip purpose, for it was found that the same model did not explain modal choices equally well in each case. Thus, the models used in this paper represent the ones that may be considered best, in the statistical sense, for each trip purpose. In the same way, the model for total travel is the model that best explains the modal choices of all travelers regardless of trip purpose. A comparison of the predictions from the total-trip model with the sum of the predictions from the models for the different trip purposes is as follows:

Item	Sum	Total
Predicted	1,183	1,183
Observed	1,184	1,184

The comparison of the combined results of the different trip purposes reveals that the predictions obtained are identical. That finding has some interesting implications for planners attempting to develop modal-choice models. In many cases, it is of interest to deal with different trip purposes separately. For example, in an investigation of staggered work hours, one might wish to consider the work trip in some detail and to ignore other purposes. Similarly, one who studies modal choice on trips to recreational sites may not wish to consider commuting trips. The results presented above have 2 interesting features in this regard. First, it is possible to isolate specific trip purposes, which may be modeled and used for predictive purposes, without being affected by the omission of the remaining trip purposes. Given the fact that planners are not always concerned with global changes, it is clearly useful to be able to separate the trip purposes. Second, it may further be implied that the models need not necessarily all be built at the same time. That may be an important consideration for the planner who may not be able to mount an effort of sufficient scale to yield models of all types simultaneously.

In more general terms, the evidence presented above indicates that it is possible to build separate disaggregate models of modal choice for different trip purposes. Moreover, all the models predict with a high degree of accuracy. For a planner concerned with subregional problems, this is an important finding. It is also demonstrated that the sum of the predictions from the separate models is equivalent to the prediction from the model for all trip purposes. The intent is not to place great emphasis on the possibility that models built separately may be combined in order to predict overall mode choices; nevertheless, an important feature of these models is that separate models may be built without impairing their joint ability to predict modal choices for all travelers.

CONCLUSIONS

A number of interesting conclusions may be drawn from the above investigation, but first it should be noted that the results are based on the empirical investigations of the Edinburgh-Glasgow area modal-split study data. As such, the results are data specific and require corroboration. It is argued that the sizes of the data sets and the consistency of the results imply generality; this should not, however, be taken as proved.

The first, and most important, conclusion is that disaggregate, behavioral, and stochastic models of modal choice are able to predict modal choices with an extremely high degree of accuracy, in some cases with prediction errors of less than 1 percent. It would have been interesting to compare these error properties with those of aggregate models based on the zone as the unit of analysis. However, the error properties of aggregate modal-choice models are not well documented, and few data sets exist that are amenable to both aggregate and disaggregate analysis. An interesting area for future research would be an explicit comparison based on the dual analysis of a single data set. In the meantime, the disaggregate results must stand alone. There can be no doubt, however, that disaggregate models predict with a high degree of accuracy.

The second conclusion is that the expected-number approach is to be preferred for deriving predictions from disaggregate models of modal choice. The classification approach is conceptually inferior, in the sense that its attempts to match predicted probabilities and observed behavior are inconsistent with the interpretation of the predicted probability. The classification approach is also an inferior predictor, as demonstrated by the fact that it produces predictions of modal-choice behavior whose prediction errors are several orders of magnitude greater than those of expected-number predictions. On both counts, the classification approach must be rejected in favor of the expected-number approach.

The third conclusion concerns the potential ability of disaggregate models to predict in circumstances that are spatially or temporally different from those under which the model was developed. The conclusion derives from the discussion of the random-division testing technique. It is argued that models should only be expected to predict well in new situations when the underlying distributions are similar to those on which the model is based. Thus, an intracity modal-choice model should not be expected to predict transatlantic mode choice; nor would a model developed for one income group necessarily predict well for another income group.

The fourth conclusion is that disaggregate models of modal choice built separately for different trip purposes yield accurate predictions for each trip purpose. In addition, the results of the separate models may be combined to yield accurate predictions of overall modal choice. That finding may have interesting implications for planners interested in subregional analyses.

In more general terms, it may be concluded that the results presented above are extremely promising and augur well for the use of disaggregate, behavioral, stochastic models in a predictive framework. A caveat is in order at this point: The predictions under consideration in this paper are point estimates. It is important that confidence intervals be developed in order that the prediction range may be assessed. It is hoped that future work will be undertaken, both to develop confidence intervals and to confirm the results of this analysis, that will lead to the use of disaggregate modeling techniques as a useful planning tool.

REFERENCES

1. Stopher, P. R., and Lisco, T. E. Modelling Travel Demand: A Disaggregate, Behavioural Approach, Issues and Implications. Proc., Transp. Res. Forum, 1970.
2. Watson, P. L. The Value of Time and Behavioural Models of Mode Choice. Univ. of Edinburgh, PhD thesis, 1972.

3. Watson, P. L. Choice of Estimation Procedure for Binary Models of Travel Choice: Some Statistical and Empirical Evidence. Northwestern Univ., Evanston, Ill., 1972.
4. Westin, R. M. Prediction From Binary Choice Models. Paper presented to Econometric Soc., Toronto, 1972.

MODAL-CHOICE AND ATTITUDE PATTERNS FOR A MEDIUM-SIZED METROPOLITAN AREA

Charles A. Hall and Vasant H. Surti, Center for Urban Transportation Studies, University of Colorado

The purposes of this study were to evaluate the factors that jointly influence the use of a public transportation system and to develop techniques of analysis and prediction that will assist the planning of future transit needs. A sampled population of employees of downtown firms in Denver was asked to complete a questionnaire. The factors that may reasonably have important influence on modal choice were analyzed. Statistical frequency distribution functions were fitted to survey data for several factors. Correlations among a number of variables were determined. The relation between certain intervals of a variable (called the conditioning variable) and data levels for other variables (called the conditioned variables) are also determined. A number of variables that would jointly have an effect on the choice of mode were used in the development of modal-choice prediction models.

•THE NEED for an adequate understanding of the factors at work in the use of multimodal transportation facilities within urban areas is particularly acute. The great and increasing degree of urbanization in the United States makes the planning of transportation facilities and the use of public transit to serve the populations of large metropolitan centers particularly urgent. Recognition of the importance of the problem by local, state, and national governments and by private organizations has led to extensive activity in collection and analysis of the present facts of urban transportation, conduct of transportation experiments under actual conditions, and investigation of alternative approaches to provision of future facilities. A major facet in understanding the needs and determining the approaches is best described as the factors influencing modal choice, the subject of this research.

BASIC ASSUMPTIONS AND OBJECTIVES

The following basic assumptions and objectives have provided the framework for this study:

1. The ultimate purposes of the research are to provide insight into the main factors that jointly influence the use of public transportation systems and to develop techniques of analysis and prediction by which the planning of future transit needs may be aided.
2. Methods of analyzing and predicting modal choice should be applicable, as far as possible, to situations differing widely in the transportation alternatives that are available or proposed and the demand that is to be satisfied.
3. It is unrealistic to expect that a high degree of precision is attainable in predicting individual behavior with respect to choice of modes and routes of travel, particularly in future or hypothetical situations. The variability and multidimensionality of choice are integral parts of the nature and heritage of Americans. However,

Publication of this paper sponsored by Committee on Social, Economic and Environmental Factors of Transportation.

recognition of the underlying factors that operate to produce roughly similar patterns of travel, independently to some extent of time and place, provides the only solid foundation for long-range planning.

4. Estimates of the effects of important factors, considered jointly, on choices among alternative modes and routes of travel must be derived and tested by application of appropriate analytical methods to a suitably large sampling of the population.

5. Although the Chicago study (1) forms a good basis, the people of Denver have different needs that can only be determined empirically.

6. The objective, then, is to determine what policies will be most effective in drawing large numbers of downtown workers away from their private automobiles and to the Denver Metro Transit (DMT) system.

BACKGROUND OF DENVER'S BUS SYSTEM

The situation in Denver closely parallels that of many American cities—the demise of public transportation. Until recently, ridership had decreased to the point where the privately owned transit firm was forced to sell the entire bus system to the city of Denver. The city has since turned the technical and management operations of the system over to a consulting firm composed of professionals in the various required areas. Ridership increased dramatically in the first 18 months of operation, and public relations activities have turned the average worker's head and at least focused his attention.

Only a satisfactory resolution of the one remaining major problem, that of having a profit ensured before a run to the suburbs is begun, stands in the way of a vastly increased ridership on the transit system. The time is surely right; ecology is a major issue, for the downtown area is covered almost daily with smog from automobile exhausts.

DATA COLLECTION

A questionnaire was prepared and distributed to employees of downtown firms. The response was excellent, partly because the questionnaires were distributed through the Downtown Denver Improvement Association and then through the appropriate personnel staffs of the individual firms.

A total of 10 CBD firms participated in the survey. Those firms were well diversified in function and size (Table 1). Adequate dispersion exists among the means for several variables to provide for objective analysis of modal choice—for example, the average annual salary as sampled was 1.64 and 1.49 ($6,600 and $6,200) for firms 02 and 10 respectively and 3.06 ($10,000) for firm 08.

MODEL FORMULATION

Model formulation proceeded from the following premises:

1. The individual trip in its entirety from origin to destination, possibly by way of specified intermediate points, is the basic unit of travel.

2. Mixed-modal trips (i.e., trips composed of segments on differing modes) are particularly important. That point of view contrasts with the usual concept of a modal split, which leads to an assignment of trips to one or another mode exclusively. The concern here is in any tendencies for travelers to switch modes either during a trip or for the total trip.

3. Various characteristics of the transportation network and characteristics of the users interact in the selection of modes and routes.

4. Estimates of the effects of important factors, considered jointly, on choices among alternative modes and routes must be derived from actual data by suitable statistical methods.

5. The analysis of factors influencing travel decisions must ultimately be based on specific information concerning decision-makers, the travel options available to them, and the decisions actually made.

6. People will usually answer questions in the proper frame of honesty; and, when the reverse is true, a rather simple analysis of the questionnaire will reveal this discrepancy so that the form can be corrected or discarded.

7. Aggregation of the data will not in itself cause any inconsistency or inaccuracies.

ANALYSES

The central purpose in the data analyses that were undertaken was to investigate basic questions concerning factors that may reasonably have important influence on modal choice.

Results of the work-trip survey made during the study are first presented in terms of characteristics of the respondents and the spectrum of kinds of trips reported. Summary values characterize the different classes of trips, such as number of trip reports, weighted frequency of occurrence, average reported travel distance, travel time, and cost. In addition, factors stated to be determinative with respect to the type of trip taken, or to be favorable or unfavorable, are listed, and the frequency of citation is given by trip class. Individual factors are grouped into more general factors—travel cost, time, convenience, comfort, and safety—and comparisons are made among trip types on the basis of relative frequency of factor citation. The reasons why persons had switched, one way or the other, between predominant use of the private automobile and predominant use of public transportation for work trips were analyzed in a similar manner. Statistical frequency distribution functions have been fitted to survey data on travel distance, travel time, and proportion of trips using public transportation. Correlations were determined among a number of the variables. In addition, the relation between certain intervals of a variable (called the conditioning variable) and data levels for other variables (called the conditioned variables) has been sought. Proportionate use of public transportation was examined in relation to selected variables taken one at a time: salary level, car driver, car ownership, distance from home to public transportation, overall travel distance, and several other factors. The proportionate use of public transportation was analyzed within the framework of multivariate regression analysis; a relatively large number of independent variables were tested. Finally, attitudes toward park-and-ride, minibus service at or near the doorstep, shuttle service from a metropolitan stadium to the downtown area, and free bus service were investigated.

Types of Trips

Nine modes of travel were selected, and each trip was categorized according to one of those modes.

Type of Trip	Category
Walk	1
Car driver	2
Car passenger	3
Bus	4
Car-pool member	5
Walk-bus (dual mode)	6
Car driver-bus (dual mode)	7
Car passenger-bus (dual mode)	8
Car-pool member-bus (dual mode)	9

Relation of Variables

A list of 29 of the variables that were included in the survey is given in Table 2. The table gives mean values for each firm surveyed and the number of applicable answers, arithmetic means, and standard deviations for the total survey.

Yearly salary levels were requested in 4 categories as follows:

Level (dollars)	Category
Under 5,000	1
5,000 to 7,500	2
7,500 to 10,000	3
Over 10,000	4

Approximately 8 percent of the respondents refused to divulge their yearly salaries. Because salary level is considered by most researchers to be correlated to personal car use, an analysis of divergent salary levels may form the basis for evaluation of several variables. Firm 08 respondents had an average salary level of 3.06 and are the only group that is significantly higher than the mean. Firms 02 and 10 each had significantly lower wages ($2,500 below the mean), and firms 03 and 04 are considerably below the norm.

The number of cars per household is also frequently considered an important variable to modal choice. In the Chicago study, the average number of cars owned per household was 1.2; in the Denver survey, the average was 1.7, a significant difference. The number of households owning various numbers of cars in the 2 cities is given below. If a weighted figure is used, far more Denverites have 2 or more cars than Chicagoans. This could explain the decline in bus ridership that took place in Denver.

Number of Cars	Chicago	Denver	Denver Weighted
0	58	62	34
1	435	478	263
2	187	556	306
3	21	139	76
4		31	17
5		7	4
6		1	1
Total	701	1,274	701

Correlation of Variables

Two distinct correlation matrices were obtained. The first included 9 variables regarding general opinions about public transit. That matrix is given in Table 3. There are no significant correlations among the distance to work (4) or the distance to DMT (11) and the other variables, except there is fairly good correlation between distance to work (4) and DMT use (12), an expected negative value (-0.26). When salary level (1) and DMT use (12) are compared, there is a -0.22 correlation, which is one of the higher values but still not a strong correlation.

The second correlation matrix is a more general review of modal choice and is given in Table 4. Several of the correlations were related by definition; others that are expected although not mandatory include distance to work (4) and distance to bus stop (11), 0.682; parking cost (15) and bus cost (16), -0.538; travel time to work (3) and distance to work (4), 0.349; number of licensed drivers (7) and number of cars (8), 0.718; total cost (34) and distance to work (4), 0.603; total cost (34) and use of a car pool (9), -0.375; and total cost (34) and log of the distance to work (30), 0.546. Of potential value to the users of the study results are several correlations including total cost (34) and use of DMT (12), -0.326, which reflects the economy of the DMT even when compared with out-of-pocket costs of only 5 cents/mile to drive; total cost (34) and frequency of DMT use (13), -0.314; total cost (34) and distance to the bus (11), 0.443; and total cost (34) and salary level squared (32), 0.590. The last item indicates a strong willingness by those having higher salaries to spend more time traveling and, conversely, by those having lower salaries to optimize costs. The salary level squared (32) is also positively correlated with travel time to work (3), 0.354; distance to work (4) in a very strong manner, 0.934; and distance to DMT (11), 0.635. The conclusion is that the lower salaried worker will optimize costs and the more highly paid worker will live farther away and will be less concerned about costs.

Table 1. Data source.

Firm	Sample Size	Total Downtown Employment	Kind of Business
01	446	1,100	Bank
02	85	–[a]	Dry goods
03	23	285	Bank
04	34	750	Bank
05	10	81	Bank[b]
06	362	3,808	Public utility
07	164	1,368	Bank
08	53	1,400	Public utility
09	44	150	Savings and loan
10	59	–[a]	Dry goods
Total	1,280	–	–

[a] Not available. [b] Combined with firm 07 for analysis.

Table 3. General preference correlation.

Variable	4	9	11	12	23	25	26	17	1
4	1.000	0.099	0.682	-0.261	-0.011	0.059	0.059	0.116	0.210
9		1.000	0.101	-0.216	0.010	0.064	0.069	-0.004	0.048
11			1.000	-0.262	-0.048	0.040	0.070	0.119	0.166
12				1.000	0.121	0.082	0.063	-0.038	0.216
23					1.000	0.200	0.178	0.133	-0.122
25						1.000	0.896	0.224	-0.023
26							1.000	0.237	0.008
17								1.000	0.069
1									1.000

Table 2. Relation of variables.

Number	Variable	Firm (mean) 01	02	03	04	06	07-05	08	09	10	Total Mean	Standard Deviation	Size
1	Salary range	2.78	1.64	2.27	2.16	2.65	2.73	3.06	2.64	1.49	2.60	1.10	1,179
2	Starting work hour	0800	0840	0810	0825	0915	0740	0750	0800	0900	0838	2.2	1,256
3	Travel time, min	39	42	43	34	50	41	43	40	42	43	17	1,184
4	Distance to work, miles	8.7	6.9	10.3	7.5	8.9	9.0	7.7	7.3	7.6	8.6	6.1	1,264
5	Can leave work on time	0.73	0.81	0.78	0.76	0.90	0.79	0.96	0.77	0.90	0.82	0.39	1,274
6	Licensed to drive	0.95	0.84	0.96	0.97	0.91	0.96	0.91	0.95	0.86	0.93	0.26	1,280
7	Number in family licensed	1.96	1.67	1.83	2.12	2.02	2.12	1.98	2.02	2.02	1.99	0.86	1,278
8	Number of cars in family	1.69	1.29	1.78	1.59	1.75	1.90	1.83	1.68	1.54	1.71	0.87	1,275
9	Car-pool member	0.24	0.24	0.17	0.35	0.21	0.33	0.30	0.18	0.15	0.24	0.43	1,280
10	Occupancy in car pool	2.4	2.4	2.3	2.3	2.7	2.3	2.7	2.0	3.0	2.3	1.0	330
11	Distance to bus stop, blocks	15.2	8.1	25.5	14.4	14.8	14.9	17.0	9.8	7.9	14.3	23.9	1,197
12	Use DMT	0.33	0.54	0.30	0.38	0.45	0.37	0.23	0.48	0.58	0.39	0.49	1,279
13	Frequency of use, days	15	18	20	16	18	17	16	16	21	17	7	502
14	Parking distance, blocks	2.9	2.1	3.3	1.2	3.0	2.7	3.2	1.7	3.5	2.8	2.6	810
15	Parking costs, dollars/month	13	19	12	13	14	16	17	17	15	14	9	809
16	Transit costs, dollars/month	12	14	16	13	14	13	12	12	16	13	6	509
17	Do park and ride	0.03	0.07	0.17	0.03	0.05	0.06	0.04	0.05	0.09	0.05	0.21	1,273
18	Percent of park-and-ride trip by bus	48	47	65	–	67	42	47	–	51	55	28	60
19	Switched modes	0.23	0.25	0.36	0.27	0.33	0.33	0.25	0.25	0.21	0.28	0.45	1,243
20	Switched mode to DMT	0.58	0.82	0.50	0.78	0.56	0.47	0.08	0.55	0.47	0.57	0.50	346
21	Aware of cost from mile-high stadium	0.76	0.73	0.87	0.76	0.87	0.77	0.78	0.93	0.64	0.80	0.40	1,203
22	Aware of free passengers on mile-high shuttle	0.61	0.60	0.41	0.59	0.65	0.62	0.57	0.84	0.48	0.62	0.49	1,177
23	Would use mile-high shuttle if stopped close to work	0.16	0.29	0.09	0.21	0.11	0.11	0.11	0.19	0.15	0.14	0.35	1,153
24	If so, how often, days	18	21	22	16	20	17	22	15	19	19	6	159
25	Would use minibus service	0.72	0.75	0.82	0.70	0.73	0.74	0.44	0.62	0.69	0.71	0.45	1,192
26	Would pay a moderate fee	0.75	0.67	0.81	0.70	0.76	0.73	0.47	0.59	0.65	0.72	0.45	1,168
27	Would use park-and-ride	0.39	0.28	0.65	0.39	0.39	0.41	0.34	0.33	0.22	0.37	0.48	1,047
28	Use free bus—same service	–	0.71	–	–	0.58	–	–	–	0.69	0.66	0.47	148
29	Use free bus—better service	–	0.85	–	–	0.86	–	–	–	0.69	0.84	0.37	146

Table 4. Correlation of values used in multistep linear regression equations.

Variable	2	3	5	4	6	7	8	9	11	12	13	17	15	16	1	30[a]	31[b]	32[c]	33[d]	34[e]
2	1.000	0.172	0.061	-0.018	-0.074	-0.020	-0.038	-0.051	-0.003	0.027	0.043	-0.053	-0.006	0.044	-0.087	-0.025	-0.011	-0.027	-0.051	0.009
3		1.000	0.040	0.349	-0.017	0.071	0.095	-0.020	0.241	0.056	0.112	0.122	-0.078	0.119	0.097	0.351	0.072	0.354	0.078	0.148
5			1.000	-0.096	-0.076	0.011	-0.034	0.040	-0.055	0.134	0.144	0.080	-0.156	0.152	-0.202	-0.114	-0.004	-0.105	-0.063	-0.172
4				1.000	0.121	0.172	0.233	0.099	0.682	-0.261	-0.228	0.084	0.108	-0.198	0.180	0.840	0.176	0.934	0.157	0.603
6					1.000	0.297	0.226	0.059	0.062	-0.220	-0.261	0.017	0.212	-0.252	0.199	0.147	0.143	0.129	0.124	0.128
7						1.000	0.718	0.102	0.140	-0.094	-0.107	0.055	0.087	-0.077	0.093	0.247	0.672	0.211	0.009	0.097
8							1.000	0.082	0.198	-0.222	-0.224	0.039	0.215	-0.211	0.126	0.299	0.938	0.269	0.592	0.222
9								1.000	0.101	-0.216	-0.283	-0.015	0.066	-0.280	0.040	0.116	0.059	0.102	0.009	-0.375
11									1.000	-0.262	-0.234	0.088	0.114	-0.212	0.134	0.533	0.172	0.635	0.130	0.443
12										1.000	0.876	0.270	-0.471	0.848	-0.208	-0.249	-0.169	-0.271	-0.219	-0.326
13											1.000	0.278	-0.538	0.967	-0.235	-0.216	-0.165	-0.241	-0.215	-0.314
17												1.000	-0.133	0.274	-0.033	0.109	0.018	0.101	0.015	-0.081
15													1.000	-0.538	0.172	0.191	0.162	0.157	0.229	0.569
16														1.000	-0.224	-0.194	-0.154	-0.214	-0.222	-0.264
1															1.000	0.210	0.145	0.190	0.116	0.177
30[a]																1.000	0.218	0.957	0.197	0.546
31[b]																	1.000	0.202	0.457	0.171
32[c]																		1.000	0.175	0.590
33[d]																			1.000	0.213
34[e]																				1.000

[a] Log of distance to work. [b] Cars squared. [c] Salary squared. [d] Car-driver ratio. [e] Total cost.

A review of the highest salary firm (08) and two low-salary firms (02 and 10) results in a verification of the general results.

Multistep Regression Analysis

The purpose of the multistep regression analysis is to assess the effects that a number of variables acting jointly have on modal choice. In particular, characteristics of persons and households are treated in conjunction with properties of the transportation choices available to them in order to clarify the complex of causes impinging on modal choice in the Denver area. The variables investigated include, as far as practicable, those that seem likely, either a priori or on the basis of substantial evidence from previous studies, to have impact on modal choice. The firms were evaluated individually and collectively. The evaluation constitutes a test, carried out with new data, of a collection of factors indicated, on grounds of inherent reasonableness or prior evidence, to be potentially useful predictors of modal choice.

The method used here is that of multivariate linear regression analysis. Only one model is formulated for each firm, with a single dependent variable (use of public transit) and several independent variables. The general model is of the form

$$Y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon$$

in which Y is the dependent variable, $x_1, \ldots, x_n$ are the independent variables, β_o, β_1, $\ldots$, β_n are the unknown coefficients of the model, and ϵ represents the net effect of contributions to the value of Y other than those specified by the first n + 1 terms on the right side.

The actual calculations were performed by a CDC 6000 computer and a stepwise regression program wherein one independent variable is added to the regression equation at each step.

The complete set of independent variables, from which multiple-regression models were developed is given in Table 5. The empirical values of the variables consist of data from the work-trip survey except as otherwise noted in the following. The definitions of most of the variables are self-explanatory; additional comments follow.

Dependent Variable—The only dependent variable studied was the use of DMT (x_{12}) as a relation between transit versus all other trips.

Independent Variables—The independent variables in group B are mathematical combinations or variations of selected group A variables and were formed to represent possible interactions among them. Each of the variables is adequately defined except x_{26}, total cost, which was derived as follows:

$$\text{Total cost} = \left\{ \frac{[(\text{distance to work}) \times (\text{cost per mile to drive car}) \times (2\ \text{trips/day}) \times (22\ \text{days} - \text{DMT use})] + (\text{parking costs})}{1\ \text{or (number in car pool)}} \right\} + \begin{array}{l}\text{transit costs}\\ \text{per month}\end{array}$$

$$x_{26} = \left\{ \frac{[(x_5)\ (0.05)\ (2)\ (22 - x_{13})] + (x_{17})}{1\ \text{or}\ x_{10}} \right\} + x_{18}$$

Results—Results are given in Table 6 for all 9 individual evaluations and for the total survey. For a majority of the analyses, 15 independent variables were entered into the equation. For the total survey, the residual variance (mean square error) of x_{12} is 0.166 with 1,264 degrees of freedom. The regression mean square is 6.382 with 15 degrees of freedom. The multiple correlation coefficient R is 0.5598, and R^2 is 0.31. Thus, 31 percent of the total sum of squared deviations of the dependent variable from its mean is accounted for by the effects of the independent variables in the equation. The standard error and F ratio to remove the variable are given in Table 6 for the total survey only. Some of the individual firms have very high R^2 values indicating little deviation from the equation.

Interpretation—A number of substantive findings emerge from the multivariate statistical analyses that have been described.

First, there is structure in the data in the sense that there is overwhelming evidence for the existence of functional relations between the dependent and independent variables. The ratio of regression to residual mean squares F is 38 for the equation estimating transit use. The expected value of F is 1 under the assumption that the independent variables have no systematic effects on the dependent variable.

The independent variables exerting major influence on the dependent variable, as measured by the individual F values, are relatively few in number. There are 15 independent variables in the regression equation of Table 6 for the total sample. There are 25 candidate independent variables.

Examination of the independent variables not included in the regression equations is of equal interest with the examination of those included. There are 8 independent variables outside the equation in every case. The following basic variables are not represented in any equation in any form: the number of riders in a car pool, frequency of public transit use, use of a park-and-ride, parking distance from work, parking cost, bus costs, mode of travel, and split between bus and car.

Exclsuion from the regression equations does not mean that the variables are totally without effect, but it does mean that the effects of the included variables are dominant. Examination of those variables would result in concurrence that they should have little or no effect on the modal choice decision.

The equation for each firm is included so one may select an equation to fit a specific set of circumstances. For instance, a successful campaign by DMT to increase ridership in dry goods companies (02 and 10) may be completely unsuccessful in other firms.

Conditioning Variable Analysis

For this analysis, the cases were separated into groups based on specified intervals of one variable, the conditioning variable. For the selected groups, computations were performed on a number of other variables that can be designated as conditioned variables. Different mean values of the conditioned variable for the different intervals of the conditioning variable would indicate a relation.

Five conditioning variables were selected: salary level, distance to work, ability to leave work on time, distance on the public transit stop, and use of DMT. This analysis was not performed for the total of all firms because the program case limits are 700 and the total sample was 1,280 cases. The analysis is based on individual firms, usually only using the larger sample sizes in order to have adequate occurrences in each of several conditioning variable intervals. Results are presented in terms of the conditioning variable.

Salary level—As salary level increases, travel time, car pool use, parking costs (with walking distance after parking down), awareness of the mile-high shuttle service, and awareness of the free passenger ride on the mile-high shuttle all increase dramatically. There is also some increase in travel time to work and number of vehicles in the family. There exists some trend toward less willingness to use the mile-high shuttle and a definite downward trend in the use of DMT. No significant correlation exists between salary level and potential use of the minibus or payment of a moderate fee for that service.

Distance to work—The distance from place of residence to work was selected as the second conditioning variable with the following intervals: <2.1, 2.1 to 5.1, 5.1 to 10.1, 10.1 to 20.1, 20.1 to 30.1, 30.1 to 50.1, and >50.1 miles (insufficient data cases exist for distances of more than 20 miles). As distance increases, travel time to work and distance to a DMT stop both increase as expected. The increased use of a car pool, higher salary, and some increased willingness to use the park-and-ride are other directly related variables. A general but not very significant reduction in use of DMT was noted, and the mile-high shuttle and minibus questions show no correlation.

Leave on time—The ability to leave work on time was opposed to not being able to leave on time and compared with 3 selected public transit modes. About twice as many (percent) use the DMT if they can and do leave work on time, probably partly as a result of having nonmanagement positions and earning less money (hence, increased ridership) and partly as a function of assurance of the availability of the public transit.

In general, almost none of those who must stay late views the park-and-ride concept as a viable option, but 8 percent of those who can leave on time would use it. Conversely, those who cannot or do not leave on time favor the minibus concept more than the others.

Distance to DMT—The distance, in blocks, to the nearest DMT stop was selected as a conditioning variable with 8 intervals: <0.1, 0.1 to 1.1, 1.1 to 2.1, 2.1 to 3.1, 3.1 to 4.1, 4.1 to 10.1, 10.1 to 20.1 and >20.1. As distance increases, the use of a car pool increases and the use of DMT increases then decreases after a few blocks. Although there is no trend in the current use of park-and-ride, there is an upward trend in willingness or desire to park-and-ride from the suburban shopping centers.

Use DMT—DMT use was divided into those who do (at least some) and those who do not, and 9 conditioned variables were evaluated. From 10 to 20 percent fewer respondents who have valid drivers' licenses use DMT, and there are about 0.3 fewer cars in the family of DMT users. Travel time to work trends upward but not significantly. There is no trend in awareness of the mile-high shuttle concepts, but there is a general trend to use mile-high shuttles and the minibus concepts (10 percent increase) if DMT is used. Park-and-ride from shopping centers seems uncorrelated.

Factors Cited as Important in Evaluation and Selection of Modes of Travel

Basic Data—Responses to 2 questions in the survey on the factors influencing modal choice are summarized in this section. The single set of factors and factor codes was set up after an initial listing was made of all the various answers given to these questions. The factors were arranged in 20 groups for convenience of coding and reference.

The 2 parts of the primary modal-choice question are repeated here to show the influence placed on the responder by a list of several potential reasons.

> Why do you take DMT? For example, safety, travel time, economy, comfort, convenience, chance to read, others need car, or bad weather.
>
> Why do you not take DMT? For example, travel time, comfort, privacy, convenience, like to drive, car needed at work, waiting, transferring cost, or exposure to weather.

As many as 4 reasons were keypunched for each respondent. If the respondent listed more than 4 reasons, a general preference code was usually keypunched. Frequently, for those who use DMT only part of the time, the factors may be both favorable and unfavorable.

The second question was answerable only if the person had switched to or from predominant use of public transportation for his work trips while maintaining his present places of employment and residence. The questionnaire heading under which replies were written read, "Main reasons for the switch." Once again, as many as 4 reasons per respondent were keypunched and are included in the analysis.

Frequencies of cases for each modal choice are shown below (from a possible 1,280 total cases).

Response	Frequency
Do not use DMT	775
Use DMT	505
Use DMT full time	293
Use DMT less than full time	212
Switch to DMT	196
Switch from DMT	150

Frequency and Aspect of Citation of Some General Travel Factors in Relation to Trip Mode—The basic data could be considered from many points of view. The approach taken is to compare various kinds of trips with respect to the frequency of citation of some of the general factors or qualities most often used to assess relative merit of travel alternatives. The factors are general preference or absence of real alternatives, travel time, cost, convenience, comfort, effort and strain of travel, danger or safety, and effects of weather. In each of the 8 citations, frequencies are given separately for the various trip classes.

The number of factor citations for each trip class or combination is divided by the number of trip reports of that class to give the relative frequency of citation (expressed as a percentage). If a factor is determinative of choice of a particular type of trip, it may also be said to be favorable or unfavorable for any specific case, so citations under the aspect of determinative can reasonably be combined when the contrast between favorable and unfavorable assessments is emphasized.

General Preference and Lack of Real Alternatives

The number of citations of the factor "general preference and lack of real alternatives" was roughly proportional to the number of trip reports in the various trip classes. There was a tendency for those who use DMT full time not to complete the rest of the questionnaire, and that may explain the 9 percent for DMT users as opposed to 6 percent for nonusers. These factors were considered as determinative of choice of travel mode rather than as favorable or unfavorable.

Travel Time

The travel time factor was cited 201 times (26 percent) by the 775 cases that do not use DMT. Of those who use DMT full time, 18 (6 percent) cited travel time as favorable. As expected, the travel time in private transportation is considered favorable and shows the willingness of many workers to drive to work in order to save a very few minutes.

Travel Cost

The travel cost factor was stated to be determinative of choice in 19 percent of all trip reports. Of non-DMT users, 53 or 7 percent cited economy as the reason for not using DMT. Many of those have free company-paid parking or are members of car pools. Of the DMT users, there were 202 citations (40 percent), of which 193 were favorable. Cost, or conversely, economy is a very strong positive factor for those using DMT.

Convenience

Convenience (Table 7) is the general travel factor most often cited in the survey. In terms of the component code groups combined under the general names, there are 421 references to convenience in the data. That total is considerably larger than the totals for any other general factor. For those who do not use DMT at all, 32 percent cited the convenience of private transportation or the inconvenience of public transportation. For those who use DMT, at least part time, the convenience factor is also very strong; 31 percent cited the convenience of public transportation.

Comfort

Comfort, specifically, was declared to be determinative of choice in 5 percent of trip reports. The general category of comfort and amenities within the vehicle was cited 175 times or about 14 percent of the time (some respondents cited more than one of these areas that would reduce the percentage slightly). Automobile users cited crowding or congestion of buses and privacy as the prime amenities. For DMT users, the ability to read, study, or work drew a large response with 43 citations (9 percent of the DMT users).

Effort and Strain of Travel

The physical and mental effort and strain of travel are covered by this general factor. Very few of the respondents (2 percent) listed this as a factor in choice of travel mode. There is a sharp contrast between driving and DMT trips with respect to this factor. A total of 25 DMT users cited this general category; that agrees with findings in the Chicago study. Of those who switched to DMT, 14 percent cited this category as the reason.

Danger or Safety

Danger or safety was rarely named as a factor by those who drive, the largest concern being safety at night (1 percent). Safety was discussed by 10 percent of those who use DMT at least part of the time. Of those who switched to DMT, 9 percent cited safety, and only 3 percent who switched from DMT listed it.

Weather

Vulnerability to weather is, to a greater or lesser extent and in various ways, a characteristic of all types of trips. Personal exposure to weather is a particular drawback for some kinds of trips, and hazardous driving is for others. Of those who do not use DMT, 50 (6 percent) cited exposure to weather; of those who use DMT, 67 (13 percent) cited bad weather driving as the undesirable aspect. Many respondents indicated that the only time they use DMT was in inclement weather.

Some Additional Specific Factors Related to Modal Choice

Detailed data on a large number of specific factors that survey respondents considered important with respect to modal choice are also analyzed. These specific factors are in addition to the more general factors treated in the preceding section. Those specific factors cited most frequently in connection with the various classes of trips are pointed out there.

Car unavailability, on the one hand, and car necessity, on the other, were very frequently cited specific factors determining nondriving and driving trips respectively. The specific factors with their frequency of citation under car unavailability are as follows:

Factor	Frequency
Car not available or not operable	24
Do not have car	30
Car needed at home	55

The specific factors under car necessity are as follows:

Factor	Frequency
Car needed or better for errands	20
Car needed for work	75
Work nonstandard hours	13

Those factors were also cited heavily in the Chicago survey. Of those who do not use DMT, 10 percent stated a need for a car at work.

"Like walking" was stated to be a desirable property of walking trips by 33 non-DMT users, and "less walking" was cited by 9. "Like to drive, do not mind driving" was a factor cited in 10 percent of the reports of driving trips in Chicago and only 3 percent in Denver. "Ecology" was cited by 36 respondents who use DMT and by 19 who switched to DMT.

"Car pool availability" was cited in a fairly large number of cases (40) along with a "ride being available" (24). The need to "pick up or discharge others," usually children, was mentioned by nearly 4 percent of the drivers.

On the unfavorable side, 10 percent of all 1,280 sampled indicated the DMT was too far away; 11 percent of those who do not use DMT cited poor DMT scheduling as the reason. An even larger number, 14 percent of the non-DMT users, disliked the waiting for late buses, and 4 percent disliked transferring. Many who have switched from DMT cited scheduling (27) or waiting (17) as the reason.

Switches Between Major Modes of Travel

There were 346 instances in which the person responded affirmatively to the question on switching major mode of travel while working in the CBD and residing at his or her present address. In 57 percent of these cases, the switch was to greater use of public transportation facilities; in the remaining 43 percent of the cases the switch was in the opposite direction. This represents a significant reversal from the situation in Chicago, where 67 percent switched away from public transit. Some further analysis is warranted and is given in Table 8 for both Chicago and Denver surveys. The specific factors from the survey are grouped into 10 classes of reasons. The frequencies of citation of reasons in each class are given separately for mode changes in the 2 directions. Frequencies are expressed as numbers of citations of factors in each class and also as a percentage of the number of switches.

Ease was cited in 53 percent of the switches to public transit in the Chicago study and in only 13 percent in the Denver study. In Denver, the largest factors influencing a switch to DMT were cost (38 percent), convenience (25 percent), and availability (18 percent). The major explanations for the difference in the percentages switching to public transit appear to be cost and convenience. For the switches from public transit, time, convenience, and availability of private transportation were predominant. The "switch-from" responses were very closely correlated between the 2 surveys.

SUMMARY AND CONCLUSIONS

The subject of the present research project, factors influencing use of the various modes of transportation in trip-making within urban areas, is highly relevant to basic decisions concerning the character of future metropolitan travel facilities. The importance of making sound decisions in the shaping of metropolitan transportation networks is now generally recognized in view of the continuing growth of population and its increasing concentration in the urban areas of the country. Transportation technology has provided a variety of feasible means and modes of urban travel including the minibus and fairly fast bus transit systems. However, intelligent evaluation of alternative possibilities in transportation planning requires that many factors other than purely technical ones be taken into account. The way a complex of transportation facilities of various modes will be used depends as much on characteristics of the population and the geographical distribution of activities as on the characteristics of the network itself. The interplay of the diverse factors that affect modal travel patterns requires for its elucidation both penetrating methods and adequate data.

The measures of effectiveness of a public transit system relate to how well and how efficiently the system serves the travel needs of the population of users. Freedom of choice among alternative ways of traveling complicates the problems of valuation and prediction; at the same time, it makes it possible to investigate empirically the factors that are important to travelers in the making of travel decisions. Research on the present project has used appropriate conceptual models in conjunction with data on travel through multimodal urban transportation systems in order to identify the main factors influencing modal choice and to quantify the effects of those factors operating jointly. Travel patterns in the large are, after all, the result of a multitude of personal choices. The best approach and one that permits both depth of causal analysis and breadth of population coverage is the statistical treatment of detailed information on a large number of individual cases. That course has been followed to the extent possible.

In conclusion, relatively few factors can explain a sizable portion of the modal choices. Although some of those are not alterable, others are and should be approached to assist in additional public transit use.

The mean number of drivers per household is 1.99 in the work-trip survey. The effect on transit use of varying the number of drivers per worker's household is certain. As the number of drivers increases, the worker is more likely to use public transportation. That is the relation to be expected because of competition for a limited number of cars.

There is also a strong effect of number of cars per household on transit use by workers. The mean number of cars owned by members of the household is 1.71 in

Table 5. Independent variables for multistep regression analysis.

Group	Variable	Symbol
A	Zip code of home address (not used in regression)	x_1
	Starting work hour	x_2
	Travel time to work, min	x_3
	Can or cannot leave work on time	x_4
	Distance to work, miles	x_5
	Are or are not licensed to drive	x_6
	Number in family licensed to drive	x_7
	Number of cars in household	x_8
	Member of car pool	x_9
	Number of riders in car pool including respondent	x_{10}
	Distance to nearest bus stop, blocks	x_{11}
	Frequency of use of DMT, round trips/month	x_{12}
	Currently use park-and-ride (dual-mode car and public transit)	x_{14}
	Location of parking for park-and-ride (not used in regression)	x_{15}
	Parking distance from work, blocks	x_{16}
	Parking cost per month, dollars	x_{17}
	Public transit cost per month, dollars	x_{18}
	Mode of travel (not used in regression)	x_{19}
	If dual mode, percentage of trip by public transit	x_{20}
	Salary level	x_{21}
B	Logarithm of distance to work = log (x_5)	x_{22}
	Square of number of cars in household = $(x_8)^2$	x_{23}
	Square of salary level = $(x_{21})^2$	x_{24}
	Car-driver ratio = (x_8/x_7)	x_{25}
	Total cost, dollars/month	x_{26}

Table 6. Multivariate regression analysis results.

Variable	Firm Values 01	02	03	04	06	07.05	08	09	10	Total	Standard Error	F Ratio
All	0.983	1.12	0.07	3.403	0.987	1.125	1.869	11.28	0.017	0.93	—	—
x_2	-0.0003	-0.0002	-0.0003	-0.0015	—	-0.00005	-0.0014	-0.102	-0.0002	-0.00007	0.00005	1.9
x_3	+0.0024	+0.0009	-0.0022	+0.0114	+0.0022	+0.0035	+0.0086	+0.0099	+0.0021	+0.0026	0.0006	15.9
x_4	+0.0073	-0.1212	+0.3289	+0.1698	+0.1521	+0.0491	+0.2088	+0.2344	+0.4531	+0.0513	0.0302	2.9
x_5	+0.0186	+0.0214	+0.2726	-0.2193	+0.0348	+0.0277	+0.1695	+0.0874	—	+0.0260	0.0064	16.4
x_6	-0.3489	-0.1356	-0.1336	—	-0.1389	-0.1560	+0.0604	-0.4726	+0.0636	-0.2467	0.0489	25.5
x_7	+0.1098	-2.066	+0.7489	-0.5487	+0.0750	-0.0911	—	-0.5783	+0.030	+0.0707	0.0321	4.9
x_8	-0.1033	+0.3527	-1.313	+0.5406	-0.3027	-0.0851	-0.2728	+0.1357	-0.181	-0.1223	0.0710	3.0
x_9	-0.3258	-0.5822	-1.081	-0.4099	-0.6669	-0.2788	-0.0354	-0.4409	-0.9492	-0.4427	0.0333	176.9
x_{11}	-0.0020	-0.0045	-0.0028	+0.0042	-0.0012	—	-0.0021	-0.0131	-0.0051	-0.0017	0.0007	6.0
x_{21}	-0.0194	+0.0908	+0.1574	+0.1570	-0.1377	-0.0512	-0.1156	-0.0342	-0.1538	-0.0373	0.0097	14.8
x_{22}	+0.6337	+0.6676	+2.456	-1.087	+0.2836	+0.0760	-0.5752	-0.7979	+1.174	+0.5020	0.1528	10.8
x_{23}	-0.0071	-0.05	+0.1583	+0.0827	+0.0492	+0.0091	+0.0276	+0.1800	+0.0443	+0.0081	0.0109	0.6
x_{24}	-0.4616	0.3463	-4.1793	+2.627	-0.4319	-0.4175	-1.471	-0.4692	-0.5813	-0.5024	0.1471	11.7
x_{25}	-0.0718	-0.2033	+0.8760	-1.887	+0.0924	-0.2135	-0.1000	-1.8784	+0.1241	-0.0204	0.0705	0.1
x_{26}	-0.0132	-0.0193	-0.0437	-0.0179	-0.0196	-0.0138	-0.0021	-0.0098	-0.0139	-0.0146	0.0012	152.9
R^2	0.27	0.43	0.93	0.58	0.43	0.31	0.49	0.59	0.68	0.31	—	—

Table 7. Frequency of citation of convenience as a factor in modal choice.

Trip Class	Trip Reports	Citation Aspect: Determinative Number	Determinative Percent	Favorable Number	Favorable Percent	Unfavorable Number	Unfavorable Percent
Do not use DMT	775	175	23	0		71	9
Use DMT	505	150	30	6	1	18	3
Use DMT part time	212	54	26	0	0	17	8
Use DMT full time	293	96	33	6	2	1	0
Switch to DMT	196	42	21	10	5	0	0
Switch from DMT	150	15	10	0	0	7	5

Table 8. Stated reasons given in Chicago and Denver work-trip surveys for switching to or from predominant use of public transit.

Reason for Switching	Chicago To Transit Number	Percent	Chicago From Transit Number	Percent	Denver To Transit Number	Percent	Denver From Transit Number	Percent
Availability of means of transportation	14	27.4	37	37.4	36	18.3	36	24.0
General preference	0	0.0	14	14.1	9	4.6	6	4.0
Cost	9	17.6	8	8.1	75	38.3	12	8.0
Time	4	7.8	37	37.4	12	6.1	35	23.3
Convenience	9	17.6	24	24.2	49	25.0	49	32.7
Comfort	6	11.8	16	16.2	4	2.0	13	8.7
Ease (less effort, strain, road congestion)	27	52.9	7	7.1	25	12.8	9	6.0
Safety, health	3	5.9	1	1.0	13	6.6	1	0.7
Environment, weather	1	2.0	5	5.0	6	3.1	3	2.0
Auxiliary activities	6	11.8	4	4.0	5	2.6	0	0.0

Denver and only 1.24 in Chicago. That factor alone explains a fair part of the past demise of Denver's public transit. There is a large decrease in transit use when 1 car is owned, and a further decrease when the number of cars owned increases to 2 and 3.

Time spent waiting for vehicles was found to have a significant effect on transit use. Many respondents indicated a long waiting time for delinquent buses.

There is a strong indication toward a willingness to use public transportation in one or more of its newer forms. More than 70 percent of the respondents indicated a willingness to use minibuses, and 37 percent indicated a willingness to use park-and-ride from outlying shopping centers. Unsolicited but welcome comments to each of those concepts were overwhelmingly favorable. They will not switch overnight, however.

As a final conclusion, all concrete results of this project, in terms of the factors that are indicated to be most influential in travel decisions, seem to be consistent with the Chicago results and with reasonable human responses to the transportation alternatives that are available.

On the basis of the research performed and the results achieved in the present project, the following recommendations are made:

1. Results of the present project show that relatively uncomplicated modal assignment models can incorporate nearly all the predictive power inherent in a fairly extensive set of independent variables. It is recommended that the variables found to be jointly most effective in the work done here be further tested in other cities or tested again in Denver after a few years of successful operation of the DMT.
2. The continuing public relations campaign of the DMT will bring results especially if accompanied by on-time service and consideration of the customers.
3. The park-and-ride concept from suburban shopping centers will meet with the same success as the Blue-Streak project in Seattle, if the buses are comfortable and express.
4. Minibus service, for free or with a moderate fee, will gain considerable ridership.

ACKNOWLEDGMENTS

The research presented in this paper is part of a project sponsored by the Urban Mass Transportation Administration. The results and views expressed are those of the authors and not necessarily those of the sponsoring agency. The authors acknowledge the assistance of the 10 CBD firms, Philip Milstein and his staff of the Downtown Denver Improvement Association, and the staff of the Denver Metro Transit.

REFERENCES

1. Bock, F. C. Factors Influencing Modal Trip Assignment. NCHRP Rept. 57, 1968.
2. Progress Report: Exclusive Busways. General Motors Corp., 1971.
3. Dixon, W. J., ed. BMD Biomedical Computer Programs. Health Sciences Computing Facility, Univ. of California, Los Angeles, 1964.
4. Surti, V. H. The Relationships of Vehicle Classification and Geometric Characteristics to Peak Period Freeway Volumes. Highway Research Record 162, 1967, pp. 86-118.

DISAGGREGATED MODAL-CHOICE MODELS OF DOWNTOWN TRIPS IN THE CHICAGO REGION

Martha F. Wigner, Chicago Area Transportation Study

Modal-choice models that combine both regional and behavioral aspects were successfully developed and calibrated for the Chicago area. The regional aspects include the coverage of trip origins throughout the entire Chicago area and the zonal nature of the data. Aspects of the models typical of disaggregated and behavioral modal-choice models are the form of the dependent variable (a dummy indicating the mode chosen) and the analytic functions used (logit and probit). Using a dummy for the dependent variable solves the problems of errors in the dependent variable and of aggregation of values of the independent variables. Probit and logit analyses restrict the value of the dependent variable suitably and are consistent with expected behavioral patterns. The independent variables chosen reflect characteristics of travelers and of the modal options available for a particular trip. These models were designed to be used both as part of the urban transportation planning package for the Chicago region and as regional planning and policy evaluation tools by themselves.

•THIS report presents the results of the calibration of modal-choice models designed to be used as part of the urban transportation planning (UTP) package for the Chicago area and also as regional planning and policy tools by themselves. The models, which to date have been calibrated only for downtown trips, combine aspects of both regional and individual behavioral modeling. Regional facets of the analysis are the coverage of trip origins throughout the entire Chicago area and the zonal nature of the data. The analytic method, the form of the dependent variable, and the choice of the variables are individual and behavioral in nature.

Because the analysis is designed to be part of a regional UTP package, the coverage of the data must include the entire region and not just a transportation corridor or some other subarea as in individual behavioral models. Therefore, zonal data are average data for the zones, and transportation system characteristics are calculated from zone centroid to zone centroid.

The individual behavioral aspects of the analysis include the choice of the analytic functions and the form of the dependent variables. The functions chosen for this analysis are logit and probit functions. They are appropriate because they restrict the values of the dependent variable between 0 and 1 and because the plotted data appear to follow the form of the curves they yield.

The form of the dependent variable in probit or logit analyses can either be the percentage of the particular modal split between an origin and destination pair (interchange) or be a dummy indicating the actual modal choice for the individual sampled trips. The former treats the modal split between 2 zones as the dependent variable; the latter treats the modal choices of individual trips. Because of the small number of trips between most origin and destination pairs in regional analysis, the modal split of an interchange is subject to large errors and is, therefore, a poor choice for the dependent variable. In addition, if interchange splits were used, they should be weighted by the

Publication of this paper sponsored by Committee on Social, Economic and Environmental Factors of Transportation.

actual number of trips for aggregation of values of the independent variables so that bias in the data is avoided. For those 2 reasons, the mode of the individual trips is the dependent variable used here.

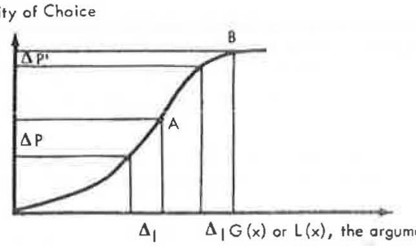
The independent variables in this analysis have been used in both regional and individual behavioral models. They reflect the characteristics that have been found to be important in individual modal-choice decisions. The variables describe the tripmaker, e.g., income, and also the particular trip, e.g., distance traveled, travel times, and travel costs.

The particular combination of regional and individual modeling used in this analysis makes possible the calibration of regional modal-choice models with increased realism and, therefore, the better ability to project future travel demands and to estimate the effects of policy and planning changes on those travel demands. Because the sample is drawn from the entire region, the results are not specific to any subarea within the Chicago area. The choice of logit and probit analyses as tools increases the realism of the model and also the accuracy of prediction to changes in the transportation system. The form of the dependent variable, a dummy indicating the choice of mode for individual trips, eliminates the problems of errors in the dependent variable and of biases in the data. The choice of the independent variables results in models that are sensitive to changes in the transportation system inasmuch as travel times and costs depend on the transportation options available between zones. Finally, the results may well be generalized to other metropolitan regions because of the behavioral nature of the analysis.

THEORY AND VARIABLES

The theoretical bases of this research are those usually found in behavioral, disaggregated modal-choice models. The distinctive features of the method are the choice of the functions used in the analysis, the form of the dependent variable, and the choice of the independent variables.

The analytic tools used in this study are probit and logit. Because possible values of the dependent variable lie between 0 and 1, a function with those limits must be found. Both the probit, or cumulative normal, and the logit functions have this characteristic. They yield S-shaped curves as shown below. The shapes of these curves are very similar although the functions themselves differ. Mathematically expressed, the probit function is



$$P = \int_{-\infty}^{G(x)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

and the logit function is

$$P = \frac{e^{L(x)}}{1 + e^{L(x)}}$$

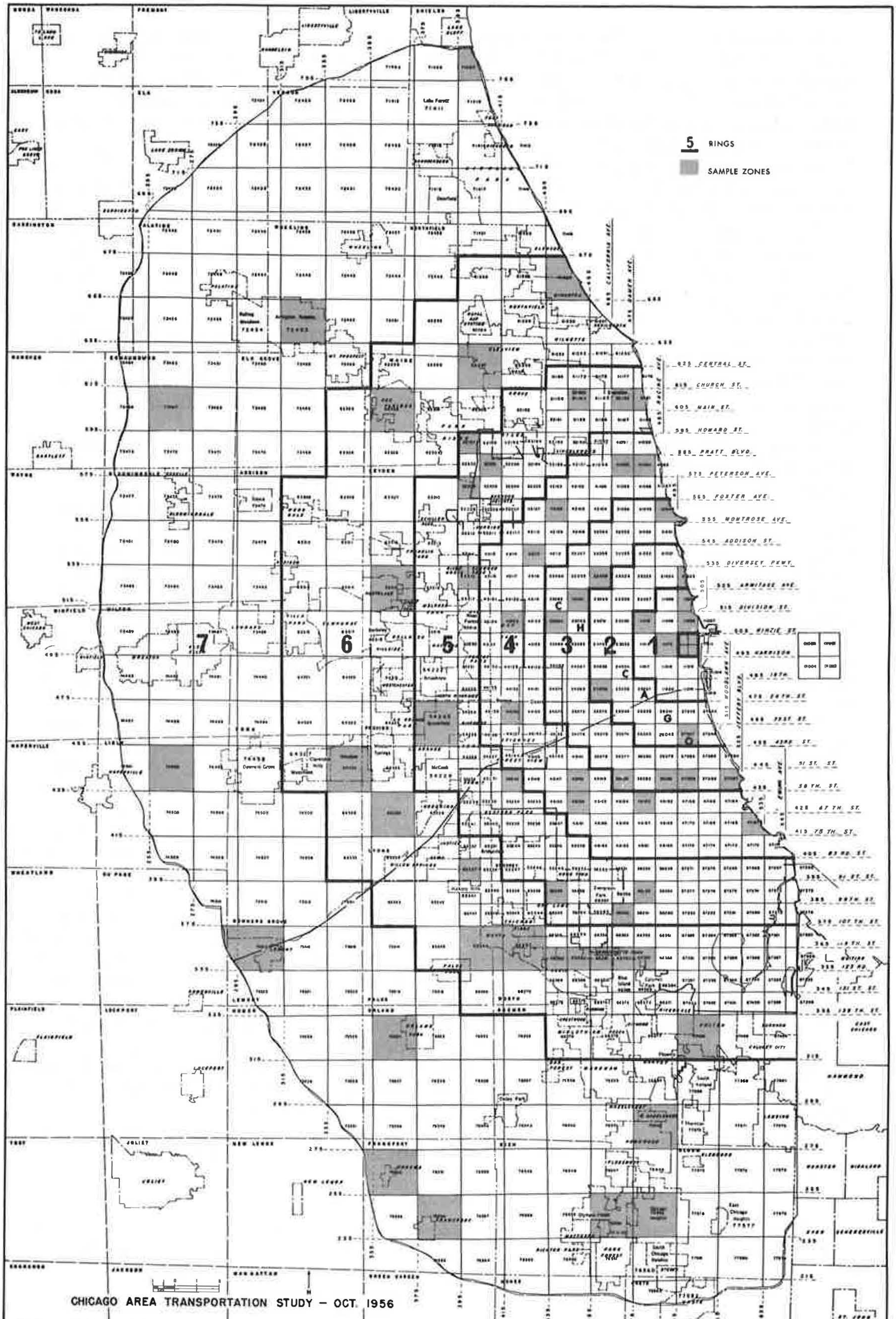
where G (x) and L (x) are linear or nonlinear functions of the independent variables, and P is the probability of modal choice.

Those functions also follow expected behavioral reactions. The effects on modal choice of given differences in travel times and costs are expected to be larger near the point of indifference between modes than at points of definite preference, where the probability of modal choice approaches 0 or 1. For example, let no preference for either mode exist at A (Fig. 1) and a definite preference exist at B. If, for example, rail travel times increase by the same amount at both A and B (Δ_1), then the expected response at A(ΔP) is greater than the expected response at B(ΔP').

The probit and logit functions also appear to be well specified, for the data follow the shapes of these particular S-curves.

For several reasons, in this analysis, as in disaggregated modal-choice analysis, the dependent variable is the modal choice of each trip instead of the modal split of all

Figure 1. Chicago Transportation Study region.



trips between 2 zones. First, the mode chosen is exactly known, while the modal split of an interchange is subject to large errors because of the small number of trips sampled for any interchange. Errors in the dependent variable result in problems of estimation that are not easily solved; the use of individual trips avoids this problem. Also their use facilitates the estimation of the mode of travel for given values of the independent variables as other factors are held constant. The unit of analysis is the trip for modal-choice analysis and the zonal interchange for modal-split analysis. Aggregation for the region results in summing trips in the first case and in summing zones in the second, unless zonal interchanges are used as weights in the summation.

Another type of weighting problem not entirely solved in this analysis occurs through the use of binary-choice models in a multimodal context. In an effort to compensate for this, we made 2 alternate assumptions concerning the structure of modal choice. First, it is assumed that each traveler makes pair-wise comparisons between modes until he decides on a mode. He considers the bus and rail modes separately as alternatives to the automobile and also to each other. Under that assumption, the sample includes all trips by either mode being analyzed. A second assumption, investigated separately, is that travelers decide first between automobile and transit modes and, then, after that initial automobile-transit decision, between the specific public transit modes.

Those 2 assumptions about traveler behavior require the definition of 3 specific and 1 combination mode: car, rail, bus, and other. The car mode was always an alternative, so it was always included in the binary-choice analysis. The rail mode includes both suburban railroad and subway (or elevated) rapid transit. Both have similar access characteristics in relation to line-haul characteristics. The stations are relatively far apart, so access costs are important. Also, line-haul travel is not in conflict with private automobile transportation, and thus congestion (as more cars enter the system) has no effect on line-haul travel. For the same reason, all public transit modes that use streets were included in the bus mode. This includes both local and express service throughout the Chicago region. The frequency of bus stops and the conflict of line-haul travel with the car mode distinguish it from the other specific transit mode. For the alternate assumption about modal-choice behavior, both public transit modes, rail and bus, were combined into an other mode.

Trips were stratified by purpose: work and nonwork. It has frequently been found that characteristics of trips associated with purpose affect the relative importance of other factors. It has been hypothesized that the repeated nature of work trips allows more thorough analysis of alternatives by workers than by others for whom the trip is infrequent. According to this reasoning, behavioral adjustment for work trips will be relatively complete because work trips continue during long periods. It is expected that behavioral adjustment will be less complete for nonwork trips because information is less complete and each trip occurs less frequently.

The specific independent variables used in the analysis are alternate times and costs, income, and trip distance. Relative travel times and costs are both explicit comparisons of modal characteristics to which travelers react. In this research, cost and time differences were used instead of ratios, an alternate method of comparison, because differences are more easily understood and several recent studies have shown that differences do a better job in explaining modal-choice behavior.

Income is included because it imposes economic limits on the amount that can be spent in efforts to save time and to increase total travel utility.

The importance to modal choice of certain factors, such as relative comfort between modes, may change with increasing distance traveled. Because comfort is hard to quantify separately, distance is assumed to be in part a proxy for it.

The constant term may be viewed as a measure of travelers' initial "bias" toward one mode or the other based on initial differential levels of comfort and convenience between modes.

DATA

The data used in this analysis are based on the CATS home interview study in 1956. The survey included a 1-in-30 sample of households in the Chicago area shown in

Figure 2. A nonrandom sample of zones was selected to keep the sample size manageable while sufficient variation in the independent variables was maintained. Sixty-one origin zones were selected from the entire region. The 6 destination zones cover the Chicago central business district (CBD). Sample sizes and average statistics for each modal choice and purpose are given in Table 1. The zones chosen are shown in Figure 2. The trip source is the actual unfactored trips from the home interview survey. Specific data items are discussed below.

Modal Choice

In modal-choice analysis, the dependent variable is 0 or 1 depending on the choice made. The automobile choice is always 1 in this analysis; the transit is 0. The transit mode is the rail mode, the bus mode, or, in the case of the other mode, the sum of the rail and bus modes.

Car Travel Times

For car travel, speeds by ring were estimated from travel times by ring to the CBD reported in the home interview survey. The car travel times in minutes were then derived from distance traveled in each ring and the estimated speed in that ring.

Public Transit Travel Times

Travel times by public transportation were estimated in minutes on the basis of line-haul time plus an estimate of access and egress times. Line-haul travel time was estimated from schedules and, if schedules were unavailable, from the 1965 CATS assignment network. Travel time by the other mode is a weighted average of the times for the specific modes. The weights are the modal splits between the bus and the rail modes. Access times were based on the walk mode unless the train station or bus stop was more than 1 mile from the centroid of the zone. In that case, access was assumed to be by car (this is consistent with the assumptions made for travel costs). Egress time downtown was always based on an assumed choice of the walk mode.

Car Travel Costs

Car travel costs were estimated in cents using distances estimated, an estimate of cost per mile (3.5 cents/mile), an estimate of the prorated cost of owning a car (32 cents/1-way trip), and half the estimated parking fee in the destination zone. The prorated cost of owning a car is the depreciation expense attributable to the trip. Parking fees in the CBD are the largest single component of car-driving costs for trips to the CBD. The values used were an interpolation of the all-day (or 8-hour) fees reported in 1948 and 1962.

Public Transit Travel Costs

Cost data were obtained from the Illinois Commerce Commission for the public transit modes. They are the 1956 fares in cents. Costs for the other mode are calculated in the same way as travel times. If the train station or bus stop was more than 1 mile from the centroid of the zone, access costs were added to the line-haul fare. Because actual mode of access to the station was not known, the access mode was assumed to be car. Capital costs were 32 cents (1-way) plus 3.5 cents/mile (parking fees were 0 at suburban stations in 1956).

Differences in Travel Costs and Times

Cost and time differences (ΔC and ΔT) were calculated by subtracting the transit travel cost or time from the automobile travel cost or time.

Income

An estimate of the average income in dollars for the families in the zone of origin was obtained from 1960 census data. Because relative income levels among zones rather than actual levels of income are the relevant income characteristic, the difference in data (1956 versus 1960) was not considered important.

Distance

Distances from the centroid of the origin zone to the centroid of the destination zone by highway were estimated in miles; the highway network existing in 1956 was used. Distance along highways was chosen rather than airline distance to reflect the differing accessibilities of areas with the same airline distance from the CBD.

ANALYSIS

Analysis of modal choice was separated by trip purpose—work and nonwork—and, for each purpose, by binary choice—car-rail, car-bus, and car-other. Graphical analysis of the data for each choice was completed first to obtain preliminary indications of the effect of individual independent variables on modal choice and to check the specification of the functions. Subsequently, various combinations of the 4 independent variables were tried in multivariate binomial probit and logit analyses. Those 2 functions yield results that are virtually identical despite differences in the values of the coefficients. Choice of technique is, therefore, a matter of taste. In the logit and probit analyses, the reliability of the coefficients and the importance of the variables were examined through the values of the standard errors of the coefficients and the value of $-2\ln\lambda$, the likelihood ratio test. The reasonableness of the coefficients was measured by using them to derive the marginal values of time and comfort and comparing them with values derived from other research. After the values of the coefficients were determined to be both statistically significant and reasonable, the relative importance of the variables among the binary choices and between trip purposes was studied, and possible reasons for differences were advanced.

Work Trips

Car-Rail Choice—Scatter diagrams of the data were drawn on normal probability paper to determine whether the postulated relations for the variables considered individually existed. Each datum point plotted consists of the percentage of car choice for about the same number of trips. If the relations were actually the S-curve specified by the probit function, then the data points would all be on a straight line. Two examples of these graphs, shown in Figures 3 and 4, illustrate the marginal effects of ΔT and ΔC on modal choice. The data points do not diverge significantly from the postulated straight line. The deviation of the actual data points from the estimated line has 3 sources. One is the true randomness in the response of individuals to a situation. The second is due to the effects of variables not included in the graphical analysis although included in later functional analysis. The third is variables omitted from the analysis.

The coefficients from logit and probit analyses for the car-rail choice for work trips are given in Table 2. Also shown are t-values, the coefficient divided by the standard error. If the t-value is 2 or greater, then, if the assumptions behind the models are correct, the coefficients differ from 0 with a probability of at least 0.98. The generally high significance (high t-values) for the coefficients of ΔC and ΔT indicates the reliability of those coefficients. The low t-values for income and distance indicate the likely lesser importance of those factors, as measured. The values of $-2\ln\lambda$ given in Table 2 confirm those results.

The signs of the coefficients of ΔC and ΔT are as expected. Increased travel cost or time for a mode results in a decreased probability of that mode being chosen. If the value of ΔC is increased by 20 cents above the average value of ΔC (\$1.10), the probability of the car choice decreases from 0.55 to 0.50. An increase in the value of ΔT by 10 min above the average value (approximately 0 min) decreases the probability of the car choice from 0.55 to 0.48. [These changes were calculated for average values of all the other variables (see Table 3)]. As ΔC and ΔT take on more extreme values, the calculated induced change in the probability of modal choice decreases because of

Table 1. Selected characteristics of samples analyzed.

Trip	Sample	Car Trips Number	Car Trips Percent	Avg Income (dollars)	Distance (miles)
Work					
Car-rail	1,314	685	52	8,400	10
Car-bus	1,518	732	48	7,800	8
Car-other	2,179	766	35	8,000	9
Nonwork					
Car-rail	756	565	74	8,000	8
Car-bus	1,037	586	57	7,500	7
Car-other	1,242	596	48	7,600	7

Figure 2. Data points and estimated probit curve for car-rail work trips to CBD—only ΔT controlled.

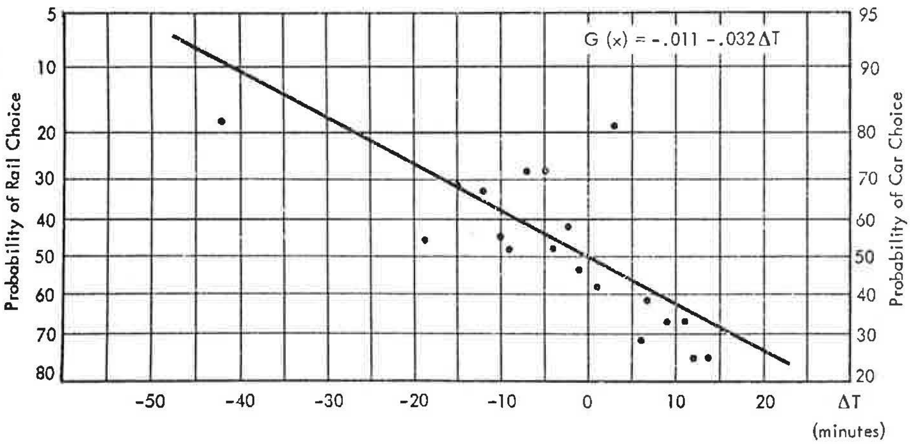


Figure 3. Data points and estimated probit curve for car-rail work trips to CBD—only ΔC controlled.

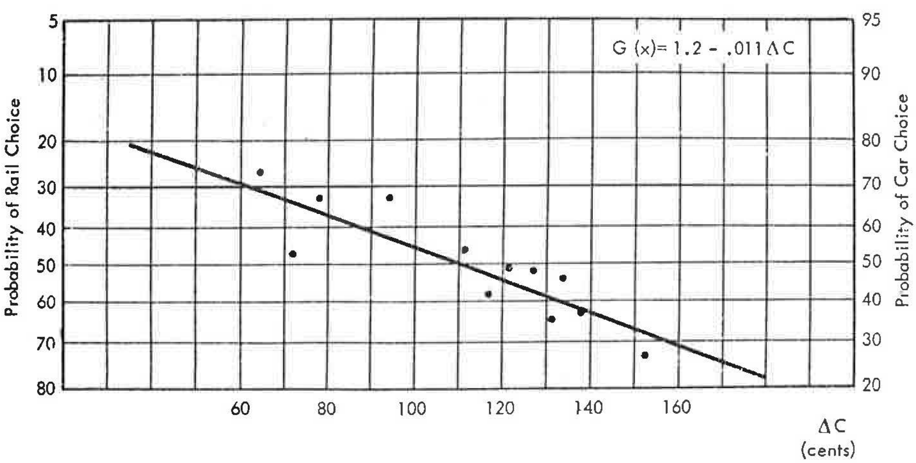


Table 2. Regression coefficients, standard errors, and t-values of car-rail work trips to CBD.

	Probit Analysis						Logit Analysis				
Equation	Constant	ΔC	ΔT	Income	Distance	$-2\ln\lambda$[a]	Constant	ΔC	ΔT	Income	Distance
Coefficient											
1	1.23	-0.011				71	1.98	-0.017			
2	-0.011		-0.032			96	-0.015		-0.052		
3	0.0076			0.55×10^{-5}		0.37	0.012			0.87×10^{-5}	
4	-0.015				0.0068	1.5	-0.023				0.011
5	0.69	-0.0062	-0.024			115	1.12	-0.010	-0.040		
6	0.75	-0.0058	-0.026	-1.2×10^{-5}		115	1.21	-0.0094	-0.042	-2.0×10^{-5}	
7	0.69	-0.0068	-0.023		0.0063	115	1.13	-0.011	-0.038		0.011
8	0.76	-0.0063	-0.024	-1.5×10^{-5}	0.0070	116	1.23	-0.010	-0.040	-2.4×10^{-5}	0.012
Standard Error											
1	0.15	0.0013					0.24	0.0021			
2	0.036		0.0033				0.058		0.0055		
3	0.14			1.6×10^{-5}			0.22			2.5×10^{-5}	
4	0.071				0.0061		0.11				0.0098
5	0.17	0.0014	0.0037				0.27	0.0023	0.0061		
6	0.19	0.0016	0.0041	1.9×10^{-5}			0.30	0.0026	0.0068	3.0×10^{-5}	
7	0.17	0.0016	0.0040		0.0070		0.27	0.0026	0.0066		0.011
8	0.19	0.0017	0.0043	1.9×10^{-5}	0.0070		0.30	0.0027	0.0071	3.0×10^{-5}	0.011
t-Value											
1	8	8					8	8			
2	0.3		9				0.3		9		
3	0.1			0.3			0.06			0.3	
4	0.2				1.1		0.2				1.1
5	4	4	6				4	4	6		
6	4	4	6	0.7			4	4	6	0.7	
7	4	4	6		0.9		4	4	6		1
8	4	4	6	0.8	1		4	4	6	0.8	1

[a]Values are the same for logit and probit analysis.

Table 3. Effect on probability of car choice of changes in value of independent variables.

	ΔC Increased 20 Cents			ΔT Increased 10 Min			Distance Increased 1 Mile			Income Increased $1,000		
Trip	From	To	Change	From	To	Change	From	To	Change	From	To	Change
Work												
Car-rail	0.55	0.50	0.05	0.55	0.48	0.07	—[a]	—[a]		—[a]	—[a]	
Car-bus	0.49	0.44	0.05	0.49	0.45	0.04	0.48	0.52	0.04	0.46	0.49	0.03
Car-other	0.35	0.30	0.05	0.36	0.32	0.04	0.32	0.34	0.02	0.34	0.36	0.02
Nonwork												
Car-rail	0.79	0.72	0.07	0.80	0.73	0.07	—[a]	—[a]		—[a]	—[a]	
Car-bus	0.59	0.51	0.08	—[a]	—[a]		0.56	0.59	0.03	0.56	0.59	0.03
Car-other	0.52	0.44	0.08	—[a]	—[a]		0.46	0.48	0.02	0.46	0.49	0.03

Note: Eq. 8 was always used.

[a]Coefficients were not significantly different from 0 at the 0.96 level.

the S-shaped curve used for analysis. The coefficients of both distance and income are small in relation to the sizes of the standard errors. Therefore, nothing can be said about them for this choice.

A check on the reasonableness of the coefficients of ΔC and ΔT was made by calculating the marginal value of time. The calculated marginal value of time is actually the amount the typical commuter to the CBD in 1956 was willing to pay to travel by a specific faster mode. The estimate for the car-rail choice for work trips is approximately \$2.30/hour. That estimate is consistent with other estimates of the marginal value of time and, therefore, provides an additional check on the sizes of the coefficients.

From the constant term and average values of income and distance traveled, a type of marginal "value of comfort" can be calculated. That is the amount people are willing to pay for the preferred mode if times and costs of the alternate modes are the same and if bus trips are ignored. The calculated marginal value of comfort includes the additional value of all factors associated with one mode as compared with the other, if trips by the third mode are ignored. In the car-rail choice, this was \$1.10 in favor of the automobile mode.

As a final check to investigate the accuracy of the specification of the analytic functions, actual and predicted probabilities of car choice were plotted as functions of G (x), the argument of the probit function (and, therefore, the estimated optimal weights for the independent variables). That is shown in Figure 5. In this case, the deviation of the actual data from the estimated curve is due only to the effect of random behavior of travelers and variables omitted from the analysis. As expected, the spread of the data points around the estimated regression line is considerably smaller when all variables are controlled than when only 1 variable is controlled (Figs. 2 and 3). There appears to be no problem of specification.

Car-Bus Choice—For the choice between car and bus modes, the estimated coefficients from probit and logit analyses for work trips are given in Table 4. The values of all the coefficients are significantly different from 0 at the 0.98 level including, in contrast with the car-rail choice, the coefficients of income and distance. That may be due to smaller differences in comfort within the bus mode in contrast with the rail mode, which includes both suburban rail and subway or elevated.

The effects of changes in the values of the independent variables are given in Table 3. Those who face the car-bus choice are less sensitive to changes in time and more sensitive to the effects of income and distance than are those who face the car-rail choice. The effect of ΔC, income, ΔT, and distance controlled is about the same. The fact that the coefficient of income is significantly different from 0 in this case indicates that the income constraint is binding.

The reasonableness of the coefficients of ΔC and ΔT was checked by a calculation of the marginal value of time for the typical commuter traveling to the CBD and faced with a choice between car and bus modes. The time value was approximately 70 cents/hour, considerably less than \$2.30 for those faced with a car-rail choice. That is consistent with an income constraint that is binding. Those who face this choice cannot afford to pay so much for their time.

The calculated marginal value of comfort for the typical commuter traveling to the CBD and faced with a choice between car and bus modes, if rail trips are ignored, is about 80 cents in favor of the automobile. That is lower than the value for those faced with a choice between car and rail modes but is consistent with the typical lower income of people who choose between automobile and bus within the same zone; they cannot afford to pay so much for greater comfort. It is also consistent with more frequent departures and more accessible stops of the bus mode, which is more convenient than the rail mode. The greater initial marginal value of comfort for people in the rail-automobile choice group in relation to those in the bus-automobile choice group is calculated to be 30 cents (\$1.10 minus \$0.80).

Car-Other Choice—If the relevant behavioral choice is between private (automobile) and public (bus and rail) modes, then the car-other choice is the relevant binary choice to analyze. The fact that income and distance were unimportant for the car-rail choice but highly significant for the car-bus choice indicates this hypothesis might be

Figure 4. Data points and estimated probit curve for car-rail work trips to CBD—optimal weights for all independent variables.

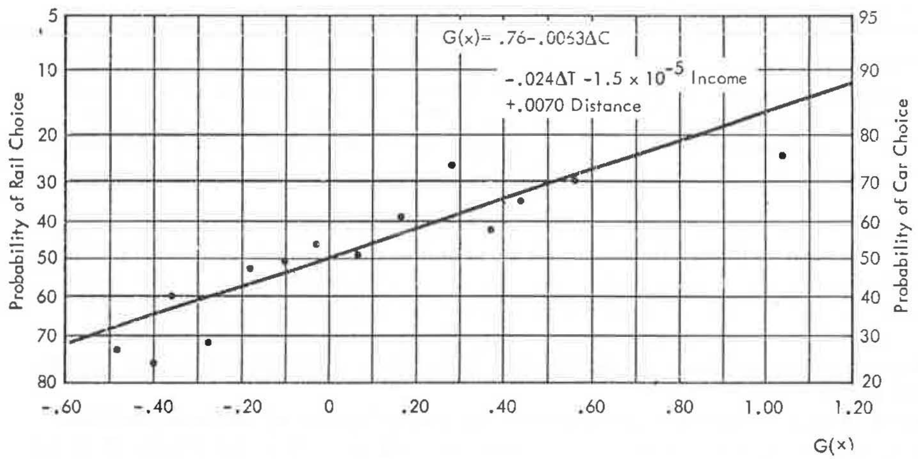


Table 4. Regression coefficients, standard errors, and t-values of car-bus work trips to CBD.

Equation	Probit Analysis Constant	ΔC	ΔT	Income	Distance	-2lnλ[a]	Logit Analysis Constant	ΔC	ΔT	Income	Distance
Coefficient											
1	-0.26	0.0020				4	-0.42	0.0033			
2	-0.39		-0.018			82	-0.62		-0.029		
3	-1.4			0.17×10^{-3}		92	-2.2			0.28×10^{-3}	
4	-0.81				0.095	194	-1.4				0.17
5	-0.56	0.0016	-0.018			84	-0.90	0.0027	-0.029		
6	-1.4	-0.00096	-0.015	0.15×10^{-3}		142	-2.2	-0.0015	-0.024	0.24×10^{-3}	
7	-0.42	-0.0056	-0.0092		0.10	234	-0.69	-0.010	-0.013		0.18
8	-0.89	-0.0063	-0.0083	0.083×10^{-3}	0.088	250	-1.4	-0.011	-0.012	0.13×10^{-3}	0.16
Standard Error											
1	0.12	0.0011					0.2	0.0018			
2	0.050		0.0021				0.082		0.0034		
3	0.14			0.018×10^{-3}			0.24			0.030×10^{-3}	
4	0.067				0.0074		0.12				0.014
5	0.13	0.0011	0.0021				0.21	0.0018	0.0034		
6	0.17	0.0012	0.0021	0.020×10^{-3}			0.27	0.0019	0.0035	0.032×10^{-3}	
7	0.13	0.0013	0.0022		0.0088		0.21	0.0022	0.0038		0.017
8	0.17	0.0013	0.0023	0.021×10^{-3}	0.0091		0.29	0.0022	0.0038	0.034×10^{-3}	0.017
t-Value											
1	2	2					2	2			
2	8		9				8		9		
3	10			9			9			9	
4	12				12		11				12
5	4	1.5	9				4	1.5	9		
6	8	0.8	7	8			8	0.8	7	7	
7	3	4	4		11		3	5	3		11
8	5	5	4	4	10		5	5	3	4	9

[a] Values are the same for logit and probit analysis.

questioned. However, t-tests on the coefficients resulting from logit and probit analyses (Table 5) and the values of $-2\ln\lambda$ all indicate significant results for this mode. The signs of the coefficients are as expected. The effects of changes in values of the independent variables tend to be intermediate between those for the specific choices (Table 3) for the typical commuter.

The calculated marginal value of time is approximately $1.15, also intermediate between the values for the 2 specific choices.

The marginal value of comfort calculated from this analysis for the typical commuter to the CBD is about 35 cents in favor of the car mode. That is less than the values for either specific choice and is due to the combination of all transit trips into one mode in the analysis of car-other trips.

Nonwork Trips

Car-Rail Choice—The results of the probit and logit analyses of the car-rail choice for nonwork trips (Table 6) indicate that the effects of distance and income could be due to chance.

The bases of the calculations of travel costs are the same for nonwork trips as for work trips. As a result, it is probable that the calculated values of ΔC are larger than the actual values. The largest component of car costs, parking fees, is probably less for nonwork trips because the stay in the CBD may be less than a full day. The difference is probably larger than the greater cost of the rail ticket because the rail ticket is bought individually and not as part of a monthly or multiride ticket. Also the number of people traveling together is greater on the average for pleasure trips than for work trips. That would result in smaller costs per person by car. The necessity for this adjustment indicates that choice for nonwork trips is probably more sensitive to changes in ΔC than implied by the calculated coefficients. The values of ΔT for nonwork trips might be modified for the same reasons as postulated for the values of ΔC. In this case, car travel time can be expected to be less because travel is not restricted primarily to peak hours, although the increased search for parking space may partially compensate for this difference. The calculated marginal time values for nonwork trips may be less reliable than that for work trips because of cost and peak—off-peak problems. The calculated value for the car-rail choice is approximately $1.10/hour for nonwork trips and is less than the $2.30/hour for work trips.

The effects on modal choice of changes in ΔC are larger for the average nonwork trip than for the average work trip; the effects of changes in ΔT are about the same (Table 3) even without the adjustments suggested above.

The marginal value of comfort if bus trips are ignored is calculated to be approximately $1.70 in favor of the automobile and is larger than the value for work trips for the same choice. It may be that those who are unfamiliar with the CBD and the transit system and those who make infrequent trips find greater certainty in traveling by car than by public transportation. The effort needed per trip to learn how to use public transit facilities is greater if the trip is infrequent than if it is frequent. Other possible reasons are that the trains run less frequently during off-peak periods and that shoppers may prefer not to carry bags to and from transit.

Car-Bus Choice—All the coefficients of the variables tend to be less significant for the car-bus choice for nonwork trips (Table 7) possibly because of data problems previously discussed. The effect on modal choice for the typical nonwork traveler facing the car-bus choice tends to be larger. The effect of ΔT was not captured for this choice (note the low t-value for the coefficient of ΔT in Eq. 8) probably because of the poor quality of the data. As a result, it was impossible to calculate the marginal value of time for this choice.

The calculated marginal value of comfort ($1.20) is less for the car-bus choice (rail trips ignored) than for the car-rail choice (bus trips ignored). As for the car-rail choice, the value is larger for nonwork trips than for work trips. The calculated marginal value of comfort for the rail mode in relation to the bus mode for nonwork trips is calculated to be 50 cents ($1.70 - $1.20), larger than the 35 cents calculated for work trips. It should be emphasized that these values assume random variations in the characteristics of the third mode.

Table 5. Regression coefficients, standard errors, and t-values of car-other work trips to CBD.

	Probit Analysis						Logit Analysis				
Equation	Constant	ΔC	ΔT	Income	Distance	$-2\ln\lambda$[a]	Constant	ΔC	ΔT	Income	Distance
Coefficient											
1	-0.019	-0.0033				12	-0.023	-0.0054			
2	-0.56		-0.017			75	-0.90		-0.027		
3	-1.1			0.086×10^{-3}		42	-1.7			0.14×10^{-3}	
4	-0.085				0.052	108	-1.4				0.086
5	-0.36	-0.0018	-0.016			79	-0.57	-0.0030	-0.026		
6	-0.87	-0.0036	-0.014	0.091×10^{-3}		120	-1.4	-0.0060	-0.024	0.15×10^{-3}	
7	-0.39	-0.0062	-0.012		0.060	199	-0.64	-0.010	-0.020		0.099
8	-0.67	-0.0068	-0.012	0.051×10^{-3}	0.054	211	-1.1	-0.011	-0.019	0.085×10^{-3}	0.089
Standard Error											
1	0.11	0.00097					0.18	0.0016			
2	0.035		0.0020				0.058		0.0032		
3	0.11			0.013×10^{-3}			0.18			0.022×10^{-3}	
4	0.054				0.0051		0.092				0.0086
5	0.12	0.00099	0.0020				0.19	0.0016	0.0033		
6	0.14	0.0010	0.0020	0.014×10^{-3}			0.23	0.0017	0.0033	0.023×10^{-3}	
7	0.12	0.0011	0.0021		0.0055		0.19	0.0018	0.0034		0.0094
8	0.15	0.0011	0.0021	0.015×10^{-3}	0.0058		0.24	0.0018	0.0034	0.024×10^{-3}	0.0097
t-Value											
1	0.2	3					0.1	3			
2	16		9				16		8		
3	10			6			9			6	
4	16				10		15				10
5	3	2	8				3	2	8		
6	6	4	7	6			6	4	7	6	
7	3	6	6		11		3	6	6		11
8	5	6	6	3	9		5	6	6	3	9

[a] Values are the same for logit and probit analysis.

Table 6. Regression coefficients, standard errors, and t-values of car-rail nonwork trips to CBD.

	Probit Analysis						Logit Analysis				
Equation	Constant	ΔC	ΔT	Income	Distance	$-2\ln\lambda$[a]	Constant	ΔC	ΔT	Income	Distance
Coefficient											
1	2.3	-0.014				53	3.8	-0.024			
2	0.66		-0.021			37	1.1		-0.058		
3	0.60			0.077×10^{-4}		0.3	0.98			0.12×10^{-4}	
4	0.67				-0.00083	0.1	1.1				-0.0013
5	2.0	-0.012	-0.015			68	3.2	-0.019	-0.034		
6	1.6	-0.011	-0.020	0.31×10^{-4}		73	2.5	-0.018	-0.042	0.70×10^{-4}	
7	1.8	-0.012	-0.019		0.013	72	3.1	-0.020	-0.037		0.029
8	1.6	-0.011	-0.020	0.23×10^{-4}	0.0052	73	2.6	-0.018	-0.042	0.54×10^{-4}	0.011
Standard Error											
1	0.24	0.0021					0.41	0.0035			
2	0.050		0.0035				0.084		0.0087		
3	0.14			0.16×10^{-4}			0.23			0.27×10^{-4}	
4	0.076				0.0067		0.12				0.011
5	0.25	0.0022	0.0037				0.45	0.0039	0.0091		
6	0.30	0.0023	0.0046	0.14×10^{-4}			0.51	0.0037	0.0083	0.25×10^{-4}	
7	0.26	0.0022	0.0043		0.0069		0.43	0.0037	0.0077		0.012
8	0.3	0.0023	0.0046	0.20×10^{-4}	0.0098		0.5	0.0038	0.0083	0.34×10^{-4}	0.016
t-Value											
1	9	7					9	7			
2	13		6				13		7		
3	4			0.5			4			0.5	
4	9				0.1		9				0.1
5	8	5	4				7	5	4		
6	5	5	4	2			5	5	5	3	
7	7	5	4		2		7	5	5		2
8	5	5	4	1	0.5		5	5	5	1.6	0.7

[a] Values are the same for logit and probit analysis.

Car-Other Choice—The coefficients in the analysis of the car-other choice (Table 8) are significantly different from 0 at the 0.98 level generally, and their signs are all as expected.

The effect on modal choice of a change in ΔC is larger for the typical nonwork traveler than for the typical commuter, the same pattern as for the 2 specific choices (Table 3). Again, as for work trips, the effect of ΔC tends to be the same for the typical pleasure trip, irrespective of the modal choice faced. The effects of distance traveled and income tend to be the same for all choices and both purposes for the typical traveler facing that choice. The effect of changes in ΔT were not captured for this choice. As a result, it was impossible to calculate the marginal value of time.

The calculated marginal value of comfort is approximately 95 cents, less than for each specific modal choice as expected from the sample selection procedure. It is larger for nonwork trips than for work trips with the same choice.

Trip Analysis Summary

The results of the analysis of modal choice of trips to the CBD are plausible and stable: The coefficients are significantly different from 0 generally, the exceptions are understandable, and the values of the likelihood ratio test ($-2\ln\lambda$) confirm these results.

The relative sizes of the coefficients between car-rail and car-bus choices and between purposes are explainable: The coefficients of ΔC and income are smaller for work than for nonwork trips, the coefficients of ΔT are larger for work than for nonwork trips, and the coefficients of distance are smaller for the car-rail choice than for the car-bus choice and larger for work trips than for nonwork trips.

The smaller sizes of the coefficients of ΔC and of income for work trips than for nonwork trips indicate demand is less price and income elastic for work trips than for nonwork trips for a given modal split. Intuitively, habit may be a more important ingredient in choice of mode for the former than for the latter in the sense that once a modal decision has been made it probably is not reconsidered unless there is a drastic change in circumstances. In contrast, the decision may be reconsidered for each nonroutine trip. In addition, even the casual traveler is likely to be aware of income and travel costs. The converse is true for factors other than cost, for example, comfort, convenience, and time savings. Those factors are more difficult to evaluate on an a priori basis, and that may explain their apparently smaller influence on nonwork trips.

The effect of distance is not significant for the car-rail choice. For the car-bus choice, the effect of distance is larger for work trips than for nonwork trips. Train and car modes have more nearly equal comfort than do bus and car modes. Therefore, as distance increases, the relative importance of comfort increases more rapidly for car-bus than for car-rail for a given modal split. Similarly, because line-haul comfort is one of the modal characteristics that is more difficult for casual travelers to investigate, the effect of distance as a proxy for comfort is less for nonwork or infrequent trips than for work trips.

The fact that the coefficients of distance and income for the car-rail choice are not significantly different from 0 at the 0.98 level was unexpected. It could be due to the inclusion of 2 submodes, the subway or elevated and the suburban railroad, in the one rail mode. That may be inappropriate data aggregation because those 2 modes have different service characteristics with respect to scheduling and comfort and serve different geographic markets that tend to be correlated with both distance and income. Therefore, the small sizes of the coefficients of income and distance may be due to an incorrect combination of submodes and not to an actual smaller effect on modal choice of distance and income.

The relative sizes of the coefficients of ΔC and ΔT were checked through calculations of the marginal value of time for the typical traveler faced with the particular choice and trip purpose. The values are reasonable, based on other similar calculations, and vary in the expected directions. The largest calculated value is for the typical commuter faced with the car-rail choice. The values, in dollars/hour, are as follows:

Table 7. Regression coefficients, standard errors, and t-values of car-bus nonwork trips to CBD.

	Probit Analysis						Logit Analysis				
Equation	Constant	ΔC	ΔT	Income	Distance	$-2\ln\lambda$[a]	Constant	ΔC	ΔT	Income	Distance
Coefficient											
1	0.28	-0.0011				0.8	0.45	-0.0018			
2	-0.057		-0.013			30	-0.092		-0.021		
3	-0.91			0.14×10^{-3}		44	-1.5			0.23×10^{-3}	
4	-0.31				0.069	58	-0.50				0.11
5	0.16	-0.0021	-0.014			32	0.25	-0.0034	-0.022		
6	-0.55	-0.0044	-0.0089	0.14×10^{-3}		64	-0.90	-0.0071	-0.015	0.22×10^{-3}	
7	0.39	-0.0086	-0.0033		0.092	89	0.64	-0.014	-0.0054		0.15
8	-0.11	-0.0091	-0.0019	0.091×10^{-3}	0.077	102	-0.17	-0.015	-0.0033	0.15×10^{-3}	0.12
Standard Error											
1	0.15	0.0014					0.24	0.0022			
2	0.056		0.0024				0.090		0.0039		
3	0.17			0.022×10^{-3}			0.28			0.036×10^{-3}	
4	0.075				0.0095		0.12				0.015
5	0.15	0.0014	0.0025				0.25	0.0023	0.0040		
6	0.20	0.0013	0.0026	0.024×10^{-3}			0.32	0.0024	0.0042	0.040×10^{-3}	
7	0.16	0.0017	0.0029		0.013		0.26	0.0027	0.0046		0.021
8	0.21	0.0017	0.0029	0.026×10^{-3}	0.013		0.34	0.0027	0.0047	0.042×10^{-3}	0.021
t-Value											
1	2	0.8					2	0.8			
2	1		5				1		5		
3	5			6			5			6	
4	4				7		4				7
5	1	1	5				1	1	5		
6	3	3	3	6			3	3	3	6	
7	2	5	1		7		2	5	1		7
8	0.5	5	0.7	4	6		0.5	5	0.7	4	6

[a] Values are the same for logit and probit analysis.

Table 8. Regression coefficients, standard errors, and t-values of car-other nonwork trips to CBD.

	Probit Analysis						Logit Analysis				
Equation	Constant	ΔC	ΔT	Income	Distance	$-2\ln\lambda$[a]	Constant	ΔC	ΔT	Income	Distance
1	0.55	-0.0056				18	0.88	-0.0090			
2	-0.14		-0.0095			13	-0.22		-0.015		
3	-0.73			0.89×10^{-4}		24	-1.2			1.4×10^{-4}	
4	-0.36				0.042	34	-0.59				0.068
5	0.41	-0.0050	-0.0081			28	0.66	-0.0081	-0.013		
6	-0.15	-0.0070	-0.0049	1.1×10^{-4}		58	-0.26	-0.011	-0.0083	1.7×10^{-4}	
7	0.50	-0.0096	-0.0014		0.062	80	0.80	-0.015	-0.0024		0.10
8	0.13	-0.010	-0.00058	0.66×10^{-4}	0.051	90	0.19	-0.016	-0.0013	1.1×10^{-4}	0.082
Standard Error											
1	0.15	0.0013					0.23	0.0021			
2	0.043		0.0026				0.069		0.0043		
3	0.14			0.18×10^{-4}			0.23			0.30×10^{-4}	
4	0.066				0.0075		0.11				0.012
5	0.15	0.0013	0.0027				0.24	0.0021	0.0043		
6	0.18	0.0014	0.0027	0.19×10^{-4}			0.30	0.0022	0.0045	0.32×10^{-4}	
7	0.15	0.0015	0.0029		0.0089		0.25	0.0024	0.0046		0.015
8	0.19	0.0015	0.0029	0.21×10^{-4}	0.0093		0.3	0.0024	0.0047	0.34×10^{-4}	0.015
t-Value											
1	4	4					4	4			
2	3		4				3		4		
3	5			5			5			5	
4	6				6		6				6
5	3	4	3				3	4	3		
6	0.8	5	1.8	5			0.9	5	1.8	5	
7	3	6	0.5		7		3	6	0.5		7
8	0.7	7	0.2	3	5		0.6	7	0.3	3	5

[a] Values are the same for logit and probit analysis.

Choice	Work	Nonwork
Car-rail	2.30	1.10
Car-bus	0.70	
Car-other	1.15	

The value of the initial modal preference is due to factors such as costs of information, reliability of the mode, frequency of departure, and, in general, the relative comfort and convenience levels of the modes. The values (in dollars) attached to these factors are as follows:

Choice	Work	Nonwork
Car-rail	1.10	1.70
Car-bus	0.80	1.20
Car-other	0.35	0.95

The calculated value of this initial preference for the specific modes changes generally in the expected directions. It is larger for the car-rail choice because the rail mode is generally less convenient than the bus mode (stations are farther apart, and trains depart less frequently). It is lower for the work purpose because public transit is generally oriented toward the commuter with respect to, for example, scheduling and costs. In addition, costs of information are lower for frequent trips than for infrequent trips. The values for nonwork trips are less stable, but again all favor the automobile mode. The low calculated value for the car-other choice in relation to the values for the specific modal choices is due to the different methods of weighting the specific modal choices and the combination other choice. For the specific choices, trips by the third mode were ignored. In the combination mode, all trips were included, but travel times and costs were weighted averages.

MODEL IMPLICATIONS AND CONCLUSIONS

The models in this study were designed for a dual purpose: to be part of the UTP package of the Chicago region and to be used as regional planning and policy evaluation tools by themselves. Those aims imposed various restrictions on the models: The trip origins had to cover the entire Chicago region, and the form of the data had to be consistent with the rest of the UTP package. Subject to those restrictions, the models were developed by using methods previously employed in behavioral and disaggregated analysis.

There are 2 basic implications. The first is that the development of an interchange modal-split model from a disaggregate modal-choice model is feasible and viable. The independent variables are sufficiently general so that the models may well be generalized to other cities and times. The second is that, given these models, the implications of certain changes in regional plans and policies can be estimated. Among these plans and policies are the introduction of a new transit facility or highway, changes in pricing policies for all modes, and changes in transfer policies and scheduling on public transit.

For example, the effect on modal choice of the introduction of a new transit facility depends on the characteristics of that system, e.g., whether it is a bus or rail line and what the transit travel times and costs are compared with those of the automobile. The effect on modal choice of a change in relative travel time can be estimated by using the results of this analysis. If a new rail line downtown were introduced and decreased rail travel time by approximately 10 min; if the origin zone were typical with respect to income, distance, and bus use to the CBD; and if there were no initial travel time advantage for either the automobile or rail mode, then rail ridership would increase by approximately 10 percentage points for work trips (from a modal split of 0.50 to one of 0.60) and by approximately 7 percentage points for nonwork trips (from 0.25 to 0.32). If, under the same conditions, the new transit route considered were a bus route and if the initial time advantage were 10 min in favor of the car mode, then a 10-min decrease in bus travel time would induce an increase in bus ridership of only 3 percentage points for work trips (from 0.54 to 0.57) and of approximately 8 percentage points for nonwork trips (from 0.34 to 0.42).

If, instead of reducing transit travel time, automobile travel time were reduced through, for example, the synchronization of traffic lights, wider car lanes, or a better or new road, trips would be diverted from all transit facilities. If the initial situation of the travelers is about what the average situation was in Chicago in 1956 and if the automobile mode had an initial advantage of approximately 10 min, then a decrease of 10 min in automobile travel time would induce a decrease in transit ridership of approximately 5 percentage points for work trips (from 0.66 to 0.61).

The effect on modal choice of changes in certain policies can also be estimated by using the results of this analysis. If a tax imposed on parking lots and garages in the CBD resulted in a flat increase in parking fees of 40 cents (or 20 cents attributable to each direction) and if the average initial cost advantage of the transit mode were $1.00, then transit ridership would increase by approximately 5 percentage points for work trips (from 0.63 to 0.68) and by approximately 8 percentage points for nonwork trips (from 0.50 to 0.58). If the tax resulted in a flat increase in parking fees of $1.00 (or 50 cents each way, starting from the same initial conditions as before), then transit ridership would increase by approximately 11 percentage points for work trips (from 0.63 to 0.74) and by approximately 19 percentage points for nonwork trips (from 0.50 to 0.69). If there were a further increase of parking fees of $1.00 (for a total of $2.00 or $1.00 each way), transit ridership would increase by only 9 percentage points for work trips (from 0.74 to 0.83) and by 15 percentage points for nonwork trips (from 0.69 to 0.84). This result—as differences (in cost) become more extreme, additional changes have a smaller effect on modal choice—is a characteristic of models using functions that yield S curves.

The effect on modal choice of a change in transfer policy can also be estimated by using the results of this analysis. If transfers within the public transit mode were facilitated by schedule changes, travel time by transit would decrease. If schedules were modified so as to decrease waiting times between the suburban railroad and the connecting distributor bus by about 5 min for the average Chicago traveler and if there were no initial time advantage for either mode, then rail ridership would increase by approximately 5 percentage points for work trips (from 0.50 to 0.55) and by approximately 3 percentage points for nonwork trips (from 0.25 to 0.28).

As suburbanization increases, it might be desirable to encourage a particular modal split. These models could be used to ascertain pricing strategies that would tend to produce the desired division. For example, if travel time by car and transit were the same, then a 50 percent modal split of work trips to the CBD would be induced by a 50-cent travel cost advantage each way for the transit mode.

These are just a few examples of types of policy and planning questions that these models could help evaluate. The response of modal choice to changes in travel times and costs depends on the initial conditions and on the extent to which the factors are varied. The reaction is strongest near the 50-50 modal-split level and decreases as the split changes in either direction. Modal-choice behavior does not change dramatically in response to relatively small changes in times and costs.

REFERENCES

1. Ergün, G. Development of a Downtown Parking Model. Highway Research Record 369, 1971, pp. 118-134.
2. Finney, D. J. Probit Analysis. Cambridge Univ. Press, 1971.
3. Lisco, T. E. The Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, PhD dissertation, 1967.
4. Rassam, P. R., Ellis, R. H., and Bennett, J. C. The n-Dimensional Logit Model: Development and Application. Highway Research Record 369, 1971, pp. 135-147.
5. Stopher, P. R. Transportation Analysis Methods. Northwestern Univ., Evanston, Ill., unpublished, 1970.

6. Stopher, P. R. A Multinomial Extension of the Binary Logit Model for Choice of Mode of Travel. Northwestern Univ., Evanston, Ill., unpublished.
7. Stopher, P. R., and Lavender, J. O. Disaggregate Behavioral Travel Demand Models: Empirical Tests of Three Hypotheses. Proc., Transp. Res. Forum 13th Ann. Mtg., Denver, Nov. 8-10, 1972.
8. Stopher, P. R., and Lisco, T. E. Modelling Travel Demand: A Disaggregate Behavioral Approach Issues and Applications. Proc., Transp. Res. Forum 11th Ann. Mtg., New Orleans, Oct. 22-24, 1970.
9. Thiel, H. A Multinomial Extension of the Linear Logit Model. Internat. Econ. Rev., Vol. 10, 1969.
10. Warner, S. L. Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice. Northwestern Univ. Press, 1962.
11. Watson, P. L. Choice of Estimation Procedure for Models of Binary Choice: Some Statistical and Empirical Evidence. Northwestern Univ., Evanston, Ill., unpublished, 1972.

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Charles V. Wootan, Texas A&M University, chairman

Committee on Social, Economic and Environmental Factors of Transportation
Harvey R. Joyner, Barton-Aschman Associates, Inc., chairman
William G. Adkins, Richard J. Bouchard, George V. Broderick, Henry W. Bruck, Jon E. Burkhardt, Leland S. Burns, John W. Fuller, William R. Helmer, Stuart L. Hill, Jonathan B. Howes, Richard Ives, Peter G. Koltnow, Paul R. Lowry, James M. Montgomery, Charles Thomas Moore, Robert L. Morris, James R. Nelson, Robinson Newcomb, Robert W. Paterson, Norman Pearson, Arthur Saltzman, Owen H. Sauerlender, John H. Suhrbier, Floyd I. Thiel, Paul Weiner, Paul F. Wendt, John Yasnowsky, Jr., Jim F. Young

Committee on Traveler Behavior and Values
Peter R. Stopher, Cornell University, chairman
David E. Bonner, Melvyn D. Cheslow, Francois X. Dedonnea, Joel Ettinger, Edward R. Fleischman, Stein Hansen, David A. Hensher, Gerald Kraft, Thomas Lisco, Robert G. McGillivray, Arnim H. Meyburg, David R. Miller, Eugene D. Perle, Shalom Reichman, Alistair Sherret, Louise Estelle Skinner, George H. Tanner, Gene R. Tyndall, Peter L. Watson, Edward Weiner, Robley Winfrey

Kenneth E. Cook and James A. Scott, Highway Research Board staff

The sponsoring committee is identified by a footnote on the first page of each report.