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

$$U^{i}_{kl/j} = d_{klo} + d_{klp}X_{ij} + e^{i}_{jkl} \qquad (1)$$

where d_{klo} is a constant, d_{klp} 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

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

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:

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

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 l (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		Some Transit	
		Adjusted Means	Standard Error	Adjusted Means	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.