The Use of Behavioral Surveys in Forecasting Transportation Requirements

JOSEPH R. STOWERS and EDMOND L. KANWIT
Respectively, Highway Engineer, and Economist, U. S. Bureau of Public Roads

This paper highlights some of the important results obtained from a nationwide survey of urban households conducted by the Survey Research Center of the University of Michigan, and shows how behavioral surveys of this type can be used to improve forecasting procedures used in urban transportation planning.

The findings are based on 824 home interviews in about 30 standard metropolitan statistical areas. Information was collected and analyzed on recent residential moves, travel habits, socio economic characteristics of households, and attitudes and preferences regarding the choice of residential location and mode of travel on the journey to work.

The critical factors influencing households' decisions on residential location and mode of travel are identified, and several conclusions are drawn about the implications of the findings for future patterns of urban development and transportation systems usage.

The paper reviews currently used methods of forecasting residential land use and trip generation, and shows how, in light of the findings of the University of Michigan study, significant improvements in these methods can be made by taking a more behavioral approach in model building.

•THIS PAPER considers the potential use of behavior surveys in highway planning and attempts to place in the perspective of current practice the findings of a recent study conducted by the Survey Research Center, University of Michigan, for the U. S. Bureau of Public Roads. The major findings of the Michigan report are presented in the first section. The forecasting implications of the behavioral approach are discussed in the second part of the paper, with special reference to residential land use and trip generation.

THE RESIDENTIAL LOCATION AND URBAN MOBILITY STUDY

The Survey Research Center study investigated intensively the attitudes and behavior of urban families, with emphasis on two types of decisions having primary bearing on urban transportation demand. Although covering a broad field of interest to city and highway planners, the study concentrated particularly on two related decision-making processes: (a) choice of location of residence, and (b) choice of mode on the journey to work.

The first report on this study was recently published by the Survey Research Center (1) and distributed to each state highway department and a number of other interested agencies and persons. Findings of the study's first phase are based on 824 personal

Paper sponsored by Committee on Economic Forecasting.

interviews from a representative cross-section of homes in standard metropolitan statistical areas. (The New York metropolitan area was excluded because of budget limitations, the area's complexities, and other considerations.)

This research effort seeks to answer questions faced by each urban area transportation study in developing and testing alternative plans for land use and transportation facilities. How will people behave under changing environmental conditions? What kind of new residential areas will be demanded, and where? How many, and what type of, trips will future residents make?

Despite the fact that the staff of every transportation study must forecast these elements, their time and budget limitations almost always preclude extensive basic research into the character of an area's urban growth and travel demand. Forecasts are necessarily based largely on trend projections of current patterns, described through massive data collection programs and analyses of highly aggregated data. This is the only course open to the transportation studies since so little precise knowledge exists about the nature of the basic factors and interrelationships involved in the decision-making processes. Forecasts based on gross descriptions of current conditions may prove reliable only in the short range.

With the foregoing state of the art in mind, the U. S. Bureau of Public Roads initiated the University of Michigan study to investigate another possible approach that might improve methods of forecasting residential land-use patterns and transportation systems usage. The specific objectives of the study were as follows:

1. To gain a better understanding of the key factors influencing choices of residential location and travel mode;
2. To measure the relative importance of the factors involved in these two areas of decision making; and
3. To gain experience with behavioral survey and research techniques to assist urban transportation studies in conducting such studies as part of their continuing planning processes.

The first report presents the findings gained from simple cross tabulations of the survey data. Additional work, now in progress, will result in a second report to be released early in 1965 covering the results of more thorough statistical analysis.

FINDINGS

The major conclusions[1] of the Residential Location and Urban Mobility study were the following:

1. The existing pattern of residential location, described by density and distance from city center, is strongly influenced by family income and by stage in the family life cycle.
2. A large proportion of potential movers would like to shift to less urbanized locations, farther away from the center of the city, into a more rural setting. Inconvenience and distance to work are unlikely to be major deterrents to outward migration.
3. The number of persons desiring to move from an apartment to a single-family house is much larger than the number interested in the opposite change. The recent boom in apartment demand reflects existing demographic and financial conditions, but not a shift in consumer preferences.
4. A widespread desire for vacation cottages or vacation homes exists and a significant proportion of families, particularly in the upper income brackets, expect to realize this ambition.
5. The number of vehicle trips per family in a 24-hr period is associated with such family characteristics as income, occupation, and age of the family head, size of the residential lot, density of neighborhood, age of the city, and number of automobiles owned. Further analysis is needed to determine the extent of independence of these

[1] This section is taken from "Summary of Findings" from the original report by the Survey Research Center, with minor revisions.

variables and their relative importance when they are taken into account simultaneously.

6. Automobile ownership and the mean annual number of miles the principal family car is driven depend on income and, to a lesser extent, on distance from city center and city age.

7. Choice of mode for the journey to work does not appear to be sensitive to cost. Most people have never even estimated the cost of driving to work; those who have report widely varying costs per mile, indicating that very little care was taken in making the estimate.

8. People would choose overwhelmingly to go to work by car rather than by common carrier if the costs and the time were the same. A frequent complaint about common carriers is that they are crowded; people prefer the freedom of movement and the convenience of travel by automobile.

FORECASTING IMPLICATIONS

General Implications

The merits of the Survey Research Center study, however, lie not so much in these 8 broad statements—for generalizations are plentiful in this field—but in the impressive array of quantified evidence behind these statements. Many ideas that had been mere conjecture before have been critically examined by the University of Michigan team. For example, the suspicion by many that the choice of mode for the journey to work was not primarily an economic decision is strongly confirmed by the report: 73 percent of all persons who drive to work have never bothered to estimate what it costs them. Again, it has long been known that there is a strong existing and potential demand for single-family housing; the present study indicates that 60 percent of families in multiple-family dwellings would prefer to live in single-family housing.

Another contribution of this survey is the fact that it cuts across cities of all types, in all parts of the country. Data collected in one metropolitan area, in an O-D study for example, apply only to that area. Such data leave unanswered the question of whether or not the relationships found are peculiar to the special socio-economic and travel characteristics of that area. With data collected in a nationwide representative sample of urban areas, it is possible to determine to what extent certain measurable characteristics of urban areas may influence deviations from the national average. Generally applicable models of behavior may be formulated and the variations among urban areas estimated.

The greatest asset of this study, however, is its comprehensive scope. Data collected for each household include not only the usual items obtained in an O-D survey, but much pertinent data usually collected by the Census of Population, as well as a wealth of information on attitudes relating to housing and modes of travel.

The move to the suburbs, to lower densities, and to increased dependence on the automobile, is a trend that will not be easily reversed. Every indication points to the assumption that this will continue to be an extremely strong force. Of all persons now living in single-family homes, the number who feel their lots are too small outnumber those who feel their lots too large by more than two to one. Until families occupy lots as large as 0.3 acre or more, those who feel they have too little land will outnumber those who feel they have too much. This implies a demand for land that would create average residential densities of about three dwelling units per net residential acre (a 100- by 130-ft lot is approximately 0.3 acre) for single-family areas.

Given a choice between a house in the suburbs on a paved street with sidewalks and lawns, or a house in the country with woods, or a field between you and the next house, 45 percent of urban residents said they would prefer the country house. The proportion tended to rise with income, with distance from downtown, and with a tendency to engage in various outdoor activities.

Even those families who stated preference for close-in living did not often give reasons which would seem to tie them strongly to the central area. Only 5 percent of these families mentioned closeness to work as a reason; even fewer mentioned social advantages of central city living. The majority, 69 percent, gave closeness to other

activities (stores, schools, etc.) as their reason for wanting to live closer in than they now do. These activities are becoming ubiquitous in modern suburbs.

A majority of respondents (52 percent) go downtown less than once a month for purposes other than work. Shopping trips outnumber other nonwork trips to the central business district by more than three to one. Some 44 percent of all recent movers stated that closeness to place of work made no difference in their choice of residential location. The proportion rises to a majority for those now living more than 10 miles from their place of work (54 percent) and those who are now 30 minutes and more from work (60 percent). Attitudes toward present and preferred housing appear to depend much more on characteristics of the house itself and qualities of the neighborhood than on locational considerations.

The study makes it obvious that most people have a poor image of central-city living. Concern about living with "nice people," social status, and neighborhood amenities undoubtedly are attitudes which are attracting people toward outlying locations.

Residential Land-Use Forecasting

The current practice among urban transportation studies in forecasting land use is to allocate anticipated total regional growth to small areas, dealing with each of perhaps half a dozen or more categories of activity separately, i.e., residential, retail, manufacturing, services, other industrial, public buildings, public open space, transportation, highways, and streets (2-6). With the exception of two or three research efforts now in progress, the allocation of population growth to new residential land is treated as if households were a single homogeneous quantity.

Typically, residential development is forecast by allocating units of growth (population or households) to available land in each small area. Factors considered in distributing growth usually include two or more of the following:

1. Either holding capacity, or the amount of land available and suitable for development;
2. Accessibility to employment sites;
3. Soil and terrain conditions (drainage, slope, soil type, etc.);
4. Level of sewer and water services; and
5. Land value.

Several weaknesses exist in this approach. Residential land deserves and requires finer analysis than this since it accounts for over half of total developed urban land and over half of all trip-ends generated. The procedure handles only the development of new residential land. It assumes implicitly that there is no change in existing built up areas. Frequently the central core area is handled by separate analysis, especially if there are plans for urban renewal or if there are areas of significant population decrease.

The new University of Michigan study demonstrates clearly the danger in treating all households alike. Present methods of forecasting do not take into account the fact that different types of households have different locational characteristics. The more important characteristics distinguishing behavior of families in residential site selection are (a) stage in the family life cycle, (b) age of the head of the household, (c) income, (d) race, and (e) education.

Also, the study shows that several of the more important factors that people consider in choosing new housing are ignored by current techniques. These include:

1. Proximity to similar types of families;
2. Type of housing available (lot size, value, number of rooms);
3. Proximity to stores, schools, etc.; and
4. Amenities of the neighborhood.

Improved methods must take a more behavioral approach to the prediction of residential growth. A significant improvement over most current methods could be made by stratifying all households into a small number of categories by stage in the family life cycle, or by income, or, preferably, by both variables. Using regression

techniques, an estimating equation could be developed for each type of household by relating the increase (or decrease) in the number of households in each small area over a recent time period, to the following variables measuring the attractiveness of each small area for that household group[2]:

1. Accessibility to households of the same stratum;
2. Accessibility to retail activity;
3. Distance to nearest school;
4. Average value of housing;
5. Holding capacity, in number of potential dwelling units, stratified by type and size of lot; and
6. Accessibility to appropriate types of employment sites.

In the forecasting process, soil and terrain conditions, general planning policies, zoning, plans for the provision of sewer and water services, and anticipated areawide demands for housing of each type should all be considered in estimating holding capacities by type of dwelling unit and size of lot.

Land-use models developed along the lines described offer several advantages. They are more reasonable conceptually and, therefore, can be used with more confidence. Since they would conform more closely with our knowledge of human behavior, the values and signs of the regression coefficients can be readily checked against hypotheses. Also, because the relationships involved in the equations are more behavioral in nature, the parameters can be expected to be more stable over time and in different metropolitan areas. Research of this type should lead toward generally applicable models of urban change.

Forecasts made with these regression equations would provide transportation planners with values of income, density, and other household characteristics for small areas, which should be helpful in estimating trip generation and modal split. Most current land-use forecasting methods will not provide these data directly.

Most importantly, these models would be sensitive to a number of key policy variables ignored by many current forecasting methods. This approach would provide another major step toward replacing the undesirable trend-based projection methods with powerful tools for comprehensive metropolitan planning.

The residential land-use forecasting method proposed poses no serious statistical complications beyond what is now typically encountered. In fact, some problems of parameter reliability would probably be simplified. One refinement that may add accuracy to the suggested procedure is worth mentioning, however.

The University of Michigan study suggests that the two variables, family life cycle and income, are most important determinants of where people tend to locate. Stratification of households into a small number of categories according to these two characteristics is definitely warranted. However, other characteristics such as family size, age, education, and occupation might also be considered.

If a large number of variables is to be considered, multivariate statistical techniques could be used to define household categories so that the behavior of each household type would be more nearly homogeneous. Component, factor, or cluster analyses could be used to aggregate categories based on the geographic correlation of household groups. Variance or regression analysis could be used to identify the most important distinguishing household characteristics for stratification of the groups. Latent class analysis has been suggested as another approach (7). This technique would define groups in a manner which would satisfy the criterion that behavior with regard to any manifest variable be independent of behavior with regard to any other manifest variable, for households within the same category. Multiple discriminant analysis might also be used.

---

[2]The regression equations for the several household types form a highly interrelated equation system. Significant bias is likely to be introduced if equation parameters are estimated by single equation least squares methods. Modern econometric methods for estimating parameters in a simultaneous equation system should be used.

Trip Generation

Many urban transportation studies base their forecasts of trip production on regression equations relating daily trips per household to one or more of three variables: automobile ownership, residential density, and distance from the central business district. The data used to develop the estimating equations are average values for traffic analysis zones, or districts, for the current year (for example, see 8). An excellent discussion of the problems involved in correlating aggregated data, and an analysis of household trip generation in Detroit and Modesto was written by Oi and Shuldiner (9).

Several problems are inherent in the approach typically used. The first of these problems is what is being measured and explained by the model, i.e., the variance between zones, not the variance between the basic units responsible for travel behavior, the household. The between-zone variance is small in relation to the between-household variance. The household is the basic observation unit for most of the data collected. Much information is lost by areal aggregation, and undoubtedly less data would be needed to develop equations with the same statistical reliability.

Relationships developed from these aggregated data are sensitive to the size of the zones and the degree of internal homogeneity achieved in drawing their boundaries. Undoubtedly this is an important factor explaining why relationships tend to differ significantly from city to city.

Relationships developed from area data will not be reliable for forecasting purposes unless the ratio of within-zone-variance to between zone-variance, with respect to all variables, can be expected to remain constant over time. On the contrary, evidence shows that newer suburban areas are less highly differentiated with respect to household characteristics than older areas in the central city (10).

Areal aggregation can obscure much of the difference in socio-economic characteristics of households and, therefore, obscure correlations that may exist between these variables and travel behavior. Cross tabulations in Residential Location and Urban Mobility show that trips per household vary with income, family size, stage in the family life cycle, occupation, as well as other variables usually considered by transportation studies such as density, automobile ownership, and distance from downtown.

Another weakness of much of the work on trip generation analysis is that too often all relationships are assumed to be linear without justification. An example would be an equation containing a positive linear regression coefficient for distance to downtown. Despite the fact the model probably overstates trip-making in existing suburbs because of the leveling off of the relationship in the outer areas, it is expected to hold valid for future fringe areas considerably further from downtown. Such assumptions become most critical when forecast values are likely to fall outside the range of current observations.

Also, it can be shown that relationships will be invalid if they are derived from aggregated data, and the basic relationships are nonlinear. As a general rule, analysis should precede aggregation rather than the reverse; the invariance of relationships with regard to area size should be verified before the analysis of aggregated data is used for forecasting purposes.

Considerable improvement might be made in trip estimation by using the household as the basic unit of analysis and by taking more care in developing relationships. Rather than assuming linearity for all independent variables, the analyst can stratify each household characteristic into several classes and treat each class as a separate dummy variable in the regression. (One class must be left out of the equation to avoid overdetermination. For a discussion of the dummy variable technique, see 11.) If a plot of the resulting regression coefficients against values of the characteristic approximates a straight line, then that household characteristic may be confidently treated as a linear variable in a subsequent equation. It should be pointed out that a simple plot of the raw data is not sufficient to verify the hypothesis of linearity, since the influence of other significant independent variables is not accounted for. An apparent linear fit in a two variable plot can, and often does, change signs, become substantially nonlinear, or become insignificant when several variables are considered simultaneously.

Preliminary analysis of the data collected in the University of Michigan study shows that transportation studies may not be focusing attention on the right variables. Table 1 gives all variables which have been found to influence trip-making significantly, in the order of their importance (the beta coefficient is a measure of the relative importance of each variable when all variables are considered simultaneously). The table also indicates which variables were approximately linear in earlier regression equations and, therefore, treated as such in the final analysis. The dependent variable is total trips made by the household per day.

Of three variables commonly analyzed in transportation studies, one is insignificant (distance from downtown) and another (density) ranks last among the other six variables. These preliminary results argue strongly for inclusion of family size, income, stage in the family life cycle, and occupation in trip generation analysis.

Trip production equations developed at the household level are easily applied to the forecasting of total zonal trip-ends. If the analyst knows the average value for each independent variable and the number of households in each category of dummy variables, he can estimate directly trip production for the zone.

For example, an estimate for a particular zone might be

$$T = Na + Nb_1\overline{X}_1 + Nb_2\overline{X}_2 + N_3b_3 + N_4b_4$$

where

T = trip-end estimate for the zone;
N = total households in the zone;
a = intercept from household regression analysis;
b_1, b_2 = regression coefficients for linear variables, from household regression analysis;
$\overline{X}_1$, $\overline{X}_2$ = average values for the zone for each of the linear variables (e.g., family size and automobiles owned);
b_3, b_4 = regression coefficients for dummy variables, from household regression analysis;
N_3, N_4 = number of households in each dummy variable category (e.g., income and family cycle classes).

The strongest argument in favor of analysis at the household level is that behavioral relationships are more likely to remain true over time and, therefore, produce more reliable forecasts. Analysis of household data, rather than area data, is needed to develop these behavioral relationships.

Thirty-five percent of the variance in trips per household was explained by the household characteristics in the analysis of the University of Michigan data. (The proportion of variance in household behavior explained by a model is substantially lower than the variance among area averages explained by models developed from area data. This is true because of the inherent daily variability in household travel habits.

TABLE 1

RANKING OF CHARACTERISTICS IN ORDER OF SIGNIFICANCE

Characteristic	Beta Coeff.	Rank
Family size (linear)	0.29	1
Automobiles owned (linear)	0.23	2
Income	0.14	3
Stage in the family life cycle	0.13	4
Occupation	0.11	5
Density of neighborhood	0.10	6
Distance	(insignificant)	7

For the errors to be directly comparable, the household model would have to be used to make estimates for the same areas for which an area model was developed.) An additional 7 percent was due to variance between cities. Further analysis showed that about one-third of the unexplained variance between cities is attributable to a few simple differences among cities, i.e., geographic region, central city population density, and age of the city. Stated another way, only 4.4 percent of the variance in household trip production is due to differences between cities which cannot be explained by the household and city characteristic variables. It does not seem unrealistic to hope that further refinements toward a general model of trip production would produce a model accurate enough to make unexplained differences among cities insignificant and thereby alleviate the need to collect trip generation data for every transportation study.

CONCLUSIONS

1. Much helpful information is to be gained in urban transportation planning by nationwide studies of the behavior of urban residents. General understanding of urban phenomena as well as forecasting methodology benefit from these efforts. Further intensive research building on the results of the University of Michigan study is definitely warranted.
2. The possibility of changing attitudes and tastes must be recognized. What will be the impact of the new towns springing up around the country? Will the image of public transit change with the introduction of modern, clean, and comfortable rapid transit systems now being planned? How much will the urban renewal programs change attitudes toward central city living? Periodic resurveys are needed to answer these and other questions.
3. In-depth behavioral surveys in individual metropolitan areas are recommended as part of the continuing planning process. Information on community attitudes will be most helpful to the planner in identifying problems, and formulating objectives.
4. Behavioral research of the type described in this paper can substantially improve our methods of forecasting and decrease the amount and cost of other large data collection programs.

REFERENCES

1. Lansing, John B., and Mueller, Eva (with Nancy Barth). Residential Location and Urban Mobility. Univ. of Michigan, Survey Research Center, June 1964.
2. Borchert, J. R., and Carroll, Donald D. Time Series Maps for the Projection of Land-Use Patterns. Highway Research Board Bull. 311, pp. 1-26, 1962.
3. Hamburg, J. R. Land Use Projection for Predicting Future Traffic. Highway Research Board Bull. 224, pp. 72-84, 1959.
4. Hansen, W. G. Land Use Forecasting for Transportation Planning. Highway Research Board Bull. 253, pp. 145-151, 1960.
5. Barnes, Charles F., Jr. Integrating Land Use and Traffic Forecasting. Highway Research Board Bull. 297, 1961.
6. Chapin, F. Stuart, Jr., and Weiss, Shirley F. Factors Influencing Land Development. Univ. of North Carolina, Aug. 1962.
7. Almendinger, V. V. Topics in Regional Growth Model: I, Penn-Jersey Transp. Study, P-J Paper No. 4, Philadelphia, April 1961.
8. Chicago Area Transportation Study. Final Rept., Vol. 1, Chicago, 1959.
9. Oi, Walter Y., and Shuldiner, Paul W. An Analysis of Urban Travel Demands. Northwestern Univ. Press, 1962.
10. Kish, Leslie. Differentiation in Metropolitan Areas. Amer. Sociological Rev., Aug. 1954.
11. Suits, Daniel. The Use of Dummy Variables in Regression Equations. Jour. Amer. Stat. Assoc., Vol. 52, Dec. 1957.