# IMPROVED TRANSPORTATION MODELING

Daniel Brand
Harvard University, Cambridge, Massachusetts

Two significant objectives for this conference were stated in remarks made at the conference: (a) to encourage those in transportation studies to devote more effort to the job of planning and less to data collection; and (b) to acquaint those attending with what is and what is not included in the 1970 census data.

## LIMITATIONS AND PROBLEM AREAS

In the past 15 years, we have spent hundreds of millions of dollars on urban transportation studies. Although our thinking about large highway and transit investment decisions has become somewhat more rigorous and systematic, I respectfully say that most of the work has not shown much. For example, a recent study at the University of Pennsylvania conducted for the Federal Highway Administration (1) concluded that the urban transportation studies found very little variation in transportation requirements (varying travel flows and needs for facilities) with varying alternative land use and transportation plans. We may ask, What is the purpose of a transportation study if not to find the best plan from among many?

There are real difficulties with the urban transportation study process. Much of the difficulty is directly traceable to the predictive models used in it. Data for predictive models, whether from the census or elsewhere, are only as useful as the models in which they are used.

The present forecasts made are quite insensitive to many of the changing values in society. Even in the more restricted area of travel forecasting, the process is insensitive to changes in technology, price, or service. In short, it is insensitive to change. It is legitimate to ask whether we are acquiring the proper kind of information on which decisions can be intelligently based.

The problem is to develop the right kinds of predictive models—perhaps different kinds of models—and then examine the data requirements for these models. We should key our information collection on current conditions and our information projection on future conditions to the kinds of decisions that must be made. (This would apply to the collection of census data or any other data. We must withhold judgment on the usefulness of census data already in hand until prior decision and modeling questions have been examined.) These include transportation investment and operating decisions involving large and small changes to the transportation network. We need to forecast the consequences of these changes. Not only are we changing our transportation system but also we are changing our values with regard to what transportation should do for high-density urban communities.

Part of the set of consequences or impacts that need to be forecast are those that concern travel or flows on the transportation system. We need to forecast travel to the required degree of disaggregation or detail of modes, links, time of day, and so on and to the required accuracy for deciding among alternative investments or operating plans. The forecasts are required for a variety of alternative plans characterized by widely varying price, speed, frequency, and convenience of the different modes in the transportation system.

In short, we need models with the required structural properties to make policy-sensitive forecasts. We need models that are relevant to our analyses, i.e., sensitive to change.

There are 2 types of models, with regard to data aggregation levels: (a) Aggregate models for which the observations are the number of trips between zones by mode, purpose, and time period when the zones have certain characteristics (averages or totals) and the transportation system has certain properties; and (b) disaggregate models for which the observations are the trips of the individual or household with certain characteristics and travel choices available to it. Which type of model is most useful, and what does this imply about census data limitations?

I suggest we must choose the latter, disaggregate models, and this puts us in direct conflict with the disclosure rules of the census. We choose the disaggregate approach, because the fundamental requirement for a demand model is that it be behavioral. It should represent decisions (in this case travel decisions) that consumers make when confronted with actual choices. We need to look at the choices that individuals make. We need to know what choices were available, which choice was made, and what characteristics that describe the trip-maker can we use to model his choice. With regard to the latter, we need data on the travelers themselves so we can isolate and model the behavior of groups exhibiting similar travel behavior in response to different transportation system characteristics. That is, the influence of different price and service characteristics of trips can be expected to vary depending on the different types of travelers and their varying utility preferences, and also depending on the nature of the final goods and services consumed or employment obtained—travel being a derived demand, not generally consumed as an end in itself. Thus, we need socioeconomic information on individual travelers at both the origins and destinations of their trips.

Does the census collect the right kinds of socioeconomic data? Yes, the right questions appear on the questionnaire: income, occupation, automobile ownership, type of dwelling unit, and employment. However, the data are sample data, and reports give only averages or totals for small areas. The data, therefore, have the same problems as socioeconomic data reported in sample home interview surveys at the zone level. The problem with such data is not the size of the area—whether zones, enumeration districts, or blocks—or the sampling rate. The problem is relating the information on individual or household trip-making behavior to its own socioeconomic characteristics and the travel choices available to it.

I would suggest, therefore, that sample zonal or block data do not answer important questions such as those that relate to changes in rates of trip-making between different socioeconomic groups or usage of modes by income groups. Nor do the data give us changes over time in these rates as needed for continuing planning. The reason is that there is great variability in socioeconomic groupings within small areas. Much information is lost in the aggregation process. Fleet and Robertson, in an excellent study (2) in Madison, Wisconsin, found 80 percent of the variance in socioeconomic data within traffic zones and only 20 percent between zones. The lost information in the tails of the distribution is often of the greatest interest. It is the information that shows where we have been and where we are headed. For example, information is lost on low-income groups about which the public sector is concerned. Who are big users of public transit? Also information is lost on high-income groups, representing presumably where we are all headed in the future. Information on these groups is generally lost in zonal averages.

Disaggregation brings with it the possibility of some real advantages and economies in data collection and simpler travel models. The use of existing census data in these models, however, has real limitations because of the strict census disclosure rules. That is, disaggregate models need information on the individual and his address (so the census data can be added to, particularly with information on the alternate travel choices). With information on the individual and his address, the name of the respondent can be traced. This is in direct violation of the (unquestioned and admirable) census disclosure rules. I suggest, therefore, that this is the primary limitation of the census travel data with respect to its use in travel forecasting. This limitation may be correctable, and such a correction will be discussed later.

There are 2 other minor problems in census travel data. The first is the reporting of the travel mode to work. The census question asks the "chief (single) means used" to get to work. Ambiguous responses are possible, particularly in the major metropolitan areas where the journey to work involves a great deal of complex travel. The "wrong" single mode could be identified, from the standpoint of the planner and modeler. Also, information is lost on the access mode of multimode trips in larger urban areas. Access to a mode appears to be very important in modeling travel demand. For example, the sensitivity of transit usage to changes in fares appears to be a function of transit service, partly as measured by access or coverage. With good service, not as many riders are lost with a fare increase. Thus, the important policy question of raising fares to cover the cost of service can be evaluated.

A second relatively minor problem in the reporting of census travel data is a possible upward bias in the number of reported work trips and the employment totals calculated by summing work trip destinations. The census question asks the "means used on the last day he worked (last week)." A simple summation of trips, including those from respondents who worked only part of the week, will likely result in an overreporting of trips. For example, a respondent who worked only 3 of 5 days will really have made only three-fifths the number of work trips on a "typical" day as the number attributed to him by a simple summation. Similarly, those who "moonlighted" or made more than 5 journeys to work will have their trips underreported.

This problem may be approximately corrected by using the information on the number of hours worked "last week" (in the census question). This involves making an assumption as to the average number of hours per day worked by the trip-maker. With this assumption, trips from trip-makers having nonstandard workweeks can be factored by the number of days worked to arrive at a more appropriate number of reported trips.

## NEEDED PROGRAMS AND RESEARCH

### The Disclosure Problem

The disclosure problem was cited previously as the primary limitation of census travel data with respect to its use in travel forecasting. Getting around the problem will require respecting the confidentiality of the data and the ability only of the Bureau of the Census to work with the data at the tally-sheet level. The bureau itself, therefore, might be requested to add the necessary modeling information to the individual trip record. This would consist primarily of adding the transportation price and service information on all available choices to the trip-maker between his origin and destination. This information, combined with the individual or household characteristics, and his trips would form one trip record. The individual addresses at the origin and destination could be destroyed, although some small area designation of the origin and destination must be preserved. Approximately 500 to 1,000 trip records, randomly selected over a region, would probably suffice to estimate a behavioral work-trip demand model.

Before the bureau is urged to do this work, however, an investigation should be made of the relative costs of such a program versus the costs of collecting and coding 500 to 1,000 similar household trip reports by conventional home interview and travel inventory means. It is quite possible that conducting new limited surveys, tailored to the particular needs of a travel forecast, would be easier, quicker, and cheaper in the long run than having the Bureau of the Census do the work.

### External Effects of Major Facilities

The external effects of major transportation facilities, such as expressways and railways, are currently receiving much attention. These are effects such as traffic noise, air and visual pollution, and barriers. Federal law since 1968 requires making public certain findings on the economic, social, and environmental impacts of highways financed with federal aid. The methodology for measuring and evaluating the impacts of external effects of transportation facilities is in its infancy.

An important program of research would utilize census data to measure the response to various measured and hypothesized external effects of expressways and rapid transit lines. Specifically, the geographic coordinates of the expressways or some other facility could be specified, and the movement history and other characteristics of residents at given distance intervals from various segments of the facilities of interest could be machine tabulated. The results would be compared with each other and with control populations located some distance away from the transportation facilities. Important data on the external effects of major facilities on traversed urban environments would be forthcoming. The tabulations would not have to be restricted to expressways, but could include all types of transportation facilities, large and small, so that comparative analyses could be made. The data could also include 1960 census data, although tabulating the latter data would be far more difficult. The additional results in terms of 1960-1970 changes in socioeconomic changes might be worth the effort, however.

### Allocation of Transit Subsidies

Several other important programs involving the use of census data in transportation planning come to mind. For example, one could investigate the question of who subsidizes and who is being subsidized by deficit-run transit service by tabulating the number of transit riders by income and (small area) location. This could be combined with information on the costs of running transit service by part of a region and the existing payments of transit deficits by parts of the region. Useful information on existing income transfers and equity solutions would be generated by such an analysis.

### Keying Information Collection to Decision-Making

In general, information collection should be keyed to the types and "value" of decisions that need to be made. Because census data represent such an exhaustive (and existing) data universe, research could draw on the existing data to determine the effectiveness of various kinds and amounts of data in reducing the chances of making wrong transportation investment decisions. The costs of such data if collected separately could be estimated. The results would be valuable in theoretical studies of the transportation planning process, leading to more effective, less costly data collection in transportation planning.

## CONCLUSION

In conclusion, we may ask ourselves again, How valuable are census data for transportation planning? As a travel universe, to stare at and provide a base at some point in time, we should, as another author suggested, test its accuracy by simply assigning it in the usual fashion. To be fair, however, we should assign it to networks by using appropriate equilibrium assignment methods, such as DODOTRANS (3).

As input to travel forecasting models, the usefulness of census travel data appears quite limited, unless we can get the Bureau of the Census itself to add the necessary information to the individual trip records. This may be the harder and more costly way to get the necessary relatively few trip records. Also, the bureau may choose to exercise its confidentiality rules by refusing small area identification of trip origins and destinations. This would eliminate the ability to use neighborhood-effect variables in the travel mode. In short, the case for the usefulness of census data in transportation planning is not at all clear. Further thought and investigations are needed.

## REFERENCES

1. Boyce, D., and Day, N. Metropolitan Planning Evaluation Strategy. Institute of Environmental Studies, Univ. of Pennsylvania, 1969.
2. Fleet, C. R., and Robertson, S. R. Trip Generation in the Transportation Planning Process. Highway Research Record 240, 1968, pp. 11-31.
3. Ruiter, E., and Manheim, M. L. DODOTRANS, A Decision-Oriented Data Organizer. Dept. of Civil Eng., Massachusetts Institute of Technology, Cambridge, various publications, 1968, 1969, 1970.