Developing a Traffic Model with a Small Sample

ROBERT G. DAVIDSON, Seminar Research Bureau, Boston College, Boston, Mass.

● FOR YEARS the Boston Metropolitan Area has accomplished too little in its efforts to create and keep up-to-date reasonably good data dealing with the pattern of daily travel origins and destinations of its residents and visitors. The first and last comprehensive O-D study was undertaken in 1945 at a time when travel patterns were abnormal because of World War II. The study must now be considered useless. In more recent years partial O-D surveys have been made, but these have consisted of Cordon Count O-D surveys and studies of industrial plan commutation patterns (such as along Route 128). These surveys cannot be classified as the type necessary to properly guide a comprehensive public transportation policy.

Nevertheless, the construction of highway and expressway facilities has progressed with moderate speed. The process has been somewhat confusing, sometimes mixed with politics, and not always in the best interests of economic development of the region. More seriously, the highway program has progressed (with Federal financial help) unilaterally—and not with the parallel development of related transportation facilities such as parking lots, transit, local streets, and frequently other highways under other governmental jurisdictions.

For those reasons, and the simple fact that those highways already built exceed estimated 1970 traffic volumes, the great need for good and up-to-date O-D data and estimates of future traffic flows is intensified. And of course it was obvious that the traffic engineering profession should have a bigger role in the decision-making process.

At this time Boston College, with Ford Foundation help, had formulated a "Seminar Research Bureau" to carry out programs of research and citizen education on Metropolitan Boston Problems. The first problem identified and investigated was urban transportation. It quickly followed that the Boston College research program would consist of investigation and use of the Gravity Model as a means, first of making (hopefully) a small contribution to the science and progression of traffic engineering, and second, making a contribution to the process of data collection and analysis in Boston. It was intended, and still is, that the Boston College Study could create a good, reliable and inexpensive traffic analysis and estimation process that would be adopted by appropriate official agencies.

The general procedure of the Boston College Study was, as follows:

1. To undertake a small, carefully designed sample home interview.
2. To use the resultant data to verify that the daily travel frequency characteristics of Boston residents are of similar magnitudes and characteristics as has been demonstrated to be the case by other O-D studies in other cities. And in this case the sample data were designed to allow establishment of relationships between social-economic characteristics of the Boston residents, and then the amount of travel taking place each day.
3. To use the sample survey data to check the general validity of the O-D distribution of daily trips as calculated by the Gravity Model formula (the model used was that developed by Alan Voorhees).
4. The remaining steps of the procedure have less to do with the use made of the sample data, but are distinctly related to the conclusions of this paper; (a) Calculating the 1959 and 1980 O-D's with the Gravity Model. The Boston College Study made 1980 calculations based on 2 alternative patterns of future regional employment—one emphasizing the downtown as a job center and the other emphasizing the suburbs near Route 128. (b) Assignment of trips to the major transportation facilities, and here it must

be noted that some six alternative assignments and systems of transport have been analyzed.

The sample home survey was designed to interview 100 families in each of ten communities. Therefore, 1,000 families were interviewed out of a total of 850,000 in the entire Boston Metropolitan Area (a 0.12 percent sample). The ten communities were selected at random from among the 84 traffic zones of the region—with the qualification that 5 of the 10 were selected intentionally from the suburbs and 5 from the inner area. The 1,000 families were also selected at random.

The interviews were carried out by senior students majoring in marketing at the Boston College School of Business Administration. About 40 students participated and were paid for time and expenses. Each student was assigned 2 homes per day for 2 or 3 days a week. The interviews were conducted during the late afternoon and evenings of October, November and early December. The questions asked of the residents were essentially the same as those asked in regular O-D surveys. The professor of the marketing course acted as consultant on interview techniques.

The sample size was small by statistical design. Because others making more extensive studies of this kind had established that the travel habits of groups of people had distinct patterns and magnitudes, and that these patterns were related to various social and economic characteristics of the people, it was necessary to verify only that these patterns and relationships also applied to the Boston region and then to modify the relationships to be somewhat more accurate and applicable. Statistically speaking, other studies had considerably narrowed the areas of uncertainness—they had defined the degrees of variability and deviation. It was statistically expected that a sample of the design carried out would allow estimation of the amount of total daily trips, and the amount of daily work trips for each community within ± 10 percent, two out of three times.

The results of the survey proved capable of doing just that. If they had not, this study would have come to an abrupt halt.

The results provided data that allowed the formulation of travel and socio-economic relationships noted, as follows:

1. Correlation between total trips per family and persons per family and cars per family. This multiple correlation has a coefficient of 0.944 and a relative variation of 8.6 percent.

2. Correlation between daily work trips and persons per family. The coefficient of correlation is 0.941 and the relative variation is 8.1 percent.

3. A good result proved to be a correlation between daily total non-work trips. The coefficient of correlation is 0.908 and the relative variation only 4.3 percent.

4. Relationships dealing with the number of more specialized types of trips did not prove to be as strong:
   (a) Shopping trips correlated with car ownership had a coefficient of 0.868 and a relative variation of 17.8 percent.
   (b) Personal business trips, business related to work, and recreation trips—as a group—related to car ownership had a coefficient of correlation of only 0.761 and a relative variation of 33 percent. The probable accuracy of prediction is not very good.
   (c) Social, civic, education and religious trips—as a group—had a coefficient of correlation of 0.819 and a relative variation of 17.7 percent when related to car ownership.

It is believed that the data from the sample were good enough and that the sample size was big enough to accomplish what was intended of it with regard to estimating travel frequency.

The results were used to calculate the 1959 and 1980 total, work and non-work travel frequency in each of the 84 zones.

The sample data were not specifically designed to guide the composition and use of the Gravity Model. It was intended only to check the distribution of trips by the model with the sample data. But after the survey was completed and the model was being checked it was noted that the value of the best exponent of the model appeared to be greater than in previous examples of its use in other cities. This could be expected to some extent because of the manner in which time and distance were measured between zones. A further explanation is the fact that a measure of "terminal time" was included in the measured travel time.

But what was disturbing was an apparent biased difference between the model and the sample data. It was noted in the investigations that a pattern suggesting that the best exponent for inner zones was different from the best exponent for outer zones when compared to the sample data.

Without going through all the details, it was finally concluded that the number of daily walk trips accounted for most of the difference and were adjusted accordingly. (Fortunately the number of walk-to-work trips had been sought out in the survey.) In the final stages the sample data established the level of the exponent in the model. The exponent settled on for work trips was 3.5 and for non-work trips was 5.0.

How accurate is the exponent? Statistically, it is not known, but the final determination of the exponent shows no bias and the distribution of trips by the model when compared to the sample data in 10 zones shows errors. In some cases large—in other cases small—as in any O-D survey estimates. But the errors show no pattern—or they have no bias. There is nothing to indicate that the errors are a function of too little data to accurately determine the travel habits from 84 different zones.

If this study were to be done again, it would be done in much the same way; but about twice as many families would be interviewed, not necessarily with the expectation of eliminating the errors that exist, but with the intent of being able to measure more accurately the errors that do occur.

But this interest is partially academic. As far as the numbers are concerned, it is felt that the normal process of aggregating and assigning the volumes tends to minimize the individual zone-to-zone O-D errors.

More impressive at this time are the differences that can occur because of various assumptions that must be made in any such study before it can rationally lead to recommendations of metropolitan transport facilities.

The influence (and related errors) of estimates of the following are important:

1. Future population distribution;
2. Future economic levels and job distribution;
3. Future car ownership;
4. The assignment procedure and its underlying assumption and/or policies regarding how easy and convenient shall travel be, what kind and how many trips shall be put on expressways, etc.;
5. The future utilization of mass transit.

There are more crucial factors that greatly affect the design of the system to be recommended. Even the simple, but not so simple to calculate, matter of future car occupancy can drastically affect the eventual design.

Who, or what O-D survey, for example, could predict that travel would increase so greatly between Cambridge and Washington, D.C.?

These kinds of influence are as important as the errors of the model. These factors show that there are many errors involved—and one of these must not be the human error of combining the very accurate with the inaccurate and expecting the perfect.

The small sample proved adequate, and even though strained, gave results that have a level of accuracy consistent with the many other inputs of a metropolitan transportation study.