Cost-Saving Techniques for Collection and Analysis of Origin-Destination Survey Data

AUSTIN E. BRANT, JR., and DANA E. LOW, Tippetts-Abbett-McCarthy-Stratton

The traditional approach in urban transportation planning requires the collection of sufficient travel data on a sample basis to permit stability at the zonal level after expansion, analysis of zonal trip generation as a function of zonal characteristics and, until recently, distribution of generated trip ends by expansion of current patterns. With the advent of models, the trip distribution process has been largely given over to simulation rather than expansion techniques. This development represents progress toward understanding more about the urban travel phenomenon and requires far less origin-destination survey data.

This paper summarizes a suggested procedure for the use of limited O-D survey data for other phases of the travel forecasting process, with attendant savings in data collection and analysis. In the area of trip generation, it is suggested that relationships be derived directly from home interview information at the household level, rather than after aggregation of O-D survey travel data and socioeconomic and land-use data to the zonal level.

A second cost-saving technique involves eliminating the traffic zone as the basic unit of analysis for all phases of the planning process. As a result, layouts of planning areas could be made that would better serve fewer functions. Other opportunities for cost savings include alternative methods of home interview sampling, use of mailed questionnaires for parts of the truck survey, and the possibility of reducing the effort usually expended in making roadside interviews by interviewing in one direction and adhering to rigid sampling techniques.

•THE preliminary steps of an urban transportation study—data collection and analysis—have been largely unchanged for 10 or 20 years. Can we do a more effective job of analysis through variations in data-gathering techniques? Do we really need as much travel data as we have traditionally accumulated for transportation planning studies? Is it necessarily true that the smaller the traffic zone size, the more refined the result obtained? These are some of the questions considered in this paper.

Comprehensiveness has not been an objective of the authors; if this paper does no more than cause some urban transportation planners to probe more deeply into the implications of their customary procedures, it will have achieved a measure of success. We must seek greater economies in pursuing urban transportation studies while we seek greater sophistication. It is the authors' contention that these are not always antithetical goals.

THE TRADITIONAL APPROACH

Since the early 1940's the home interview type of origin-destination survey has been used to provide the basic travel data for urban transportation planning studies. Methods

Paper sponsored by Committee on Origin and Destination and presented at the 46th Annual Meeting.

of conducting O-D surveys of this type have become a part of urban planning "tradition"—in Webster's definition, "something handed down from the past." Measured in years and months this "past" is not so long, but in terms of the number of cities in which these surveys have been made (several hundred in the United States alone during the past two decades), number of people affected (millions), and total cost (considerable), this "past" is extensive.

The three basic ingredients of a typical O-D survey of the home interview type are the dwelling unit survey, the truck and taxi survey and the roadside interview survey. In the dwelling unit survey, occupants of a carefully controlled sampling of dwelling places are interviewed to obtain detailed data about personal travel in the urban area under study. Considerable information about the occupants themselves is also recorded. In the roadside interview survey, selected drivers entering or leaving the study area are questioned about the trip that they are making into or out of the area. When the travel data obtained in each of these sample interview surveys are properly expanded and combined, a complete description of current travel in the study area is obtained.

Traffic zones are traditionally used as the basic areal units for summarizing and analyzing the vast amounts of data obtained in the interview surveys. A long list of criteria governing the manner in which the study area should be subdivided into traffic zones could be prepared, but basically they are all aimed either at minimizing the distortions that arise from aggregation of data or at conforming to areal definitions for which other data are or will be available.

Expansion of sample data from a full-scale O-D survey yields statistically reliable measures of both current trip generation and current trip distribution. Forecasting future trip generation usually involves development of trip production and attraction equations by means of multiple regression analysis, in which zonal trip ends within the study area determined from the O-D survey are related to various socioeconomic and land-use parameters of the zones. Separate equations are usually derived for various trip purposes and sometimes for different periods of the day. Future zonal estimates of the independent socioeconomic and land-use variables are introduced into trip generation equations to determine future zonal trip productions and attractions.

TRIP DISTRIBUTION

Travel patterns are the result of an immensely complicated interweaving of forces brought about by the spatial separation of people from the activities, places and other people that play a part in their lives. Growth factor procedures for forecasting trip distribution have a number of well-known limitations which are really all reflections of the fact that they "recognize" these forces without "explaining" them.

Philosophically at least, the use of models represents a considerable advance over the older techniques. Models which incorporate general theories and seek to simulate complex phenomena lead us to understanding of these phenomena. From this understanding we can hope to use the general theories with no more than a simple calibration process to account for special or unusual characteristics of the area in which they are to be applied. In transportation planning, models have been most successfully applied to trip distribution. The success of this application has led the authors to the following conclusions:

> Proposition 1—We should seek to emulate the philosophical strength of the trip distribution models in all areas of transportation planning. That is, we should try to develop procedures that will "explain," not simply "recognize" the significance of diverse factors that influence human travel behavior.
>
> Proposition 2—When we have accomplished the above, we need only calibrate to account for differences from the norm. We are not required to derive a basic theory for each particular area which we are planning. It follows that considerably less travel data should be sufficient, and therefore considerable cost savings should be possible.

With the gravity distribution model, for example, we require a set of friction factor curves. Current practice is to develop these to "fit" a particular area by trial and error procedures, but the first trial is generally to use available curves developed for some other area with similar characteristics. And when the model is calibrated the final curves are really not so very different, in general appearance anyway, from the initial curves. It has been shown that a very small random home interview sample will provide sufficient data for calibration of a gravity trip distribution model in a small urban area (1). The authors believe that further research will point to the same conclusion for larger urban areas.

At the heart of the approach suggested in this paper is the belief that a small sample will be sufficient foundation for urban travel forecasting if structured and analyzed on some other basis than traffic zones. No suggestions are made here for reducing the quantity or quality of socioeconomic and land-use data customarily considered necessary for urban transportation planning studies. As a matter of fact, many of the proposed procedures are dependent upon having such data in considerable detail and at a high level of reliability.

TRIP GENERATION

Most urban transportation studies have analyzed trip generation using multiple regression techniques on data aggregated by zones or districts. Generation equations derived in this manner usually prove to be quite reliable, and seem to explain trip production and attraction to a reasonable degree in the area for which they are developed. Assuming that these equations really do get to the causes of trip-making in the particular area for which they were developed, does the fact that equations for the same types of trips developed in different urban areas seldom bear much resemblance to one another force us to conclude that people are really that much different, insofar as their travel habits are concerned, from area to area? The authors suggest that the process of zonal or district aggregation of data is actually wasting much information collected in the interviews, and that this aggregation procedure in itself may be the cause of many of the variations between different areas. It is suggested in this paper that the household would make a better analysis unit for this purpose.

The authors believe that fewer data will be required and better results will be obtained if portions of the trip generation analysis are carried out at the household level, with each home interview representing an observation. Using this procedure, the sample size would depend primarily on the range of social and economic stratifications of the population in the area, and would be set to obtain adequate household and travel data within each stratification, without too much regard for the geographic extent of the area or its population. The same small random home interview sampling required to provide data for calibration of a gravity trip distribution model will very likely yield enough information for this type of generation analysis.

It might be well to introduce a word of caution at this point. Although the authors feel that more meaningful results are potentially available from the suggested type of analysis, the statistical measures of accuracy normally used to evaluate how good an equation is may not look as favorable. Much of the variance among samples is dampened as a result of aggregating data to the zonal level. Of course, much of the essential meaning may have been lost, too, even though the statistical correlation of zonal averages looks better.

Home-based trip production equations would be developed using home interview data pertaining to home-based trips and household characteristics. Multiple regression analyses would be carried out on an interview basis, with the number of home-based trips (perhaps stratified by purpose) per household taken as the dependent variable, and various socioeconomic or land-use characteristics of the household as independent variables. If the resulting equations were linear, they could be used directly in forecasting of home-based trip productions on a zonal basis by simply entering zonal forecast averages of the independent variables and multiplying by the forecast number of households in each zone. If nonlinear variables were involved, separate equations could be developed for different stratifications of the nonlinear variables, in which case some special treatment would be required to insure that the actual distribution of the nonlinear variables were properly incorporated into the final zonal forecasts. Or values taken by an independent variable could be stratified into several classes and each class made a dummy variable in the regression equation. Using this technique, the dummy takes on values of either 1 or zero for each observation in the regression analysis depending on whether or not the variable falls within the dummy class.

Trips with one end at home are generally considered to be produced at home, regardless of whether they are to or from home. There is great logic in so relating trip production to the people who make the trips, and to the socioeconomic and land-use characteristics of the household which are indicative of the reasons behind their desire and ability to make trips. When it comes to non-home-based trips, however, the logic of the traditional mode of analysis diminishes. Because neither end is at home, we customarily take the origin end to be the production end and seek out a relationship between non-home-based trips produced and circumstances at the production end. It is suggested that it would be more logical to consider these trips to be generated at home. Exactly as with home-based trips, it is the complex combination of circumstances that we can measure only at home that fashions the travel desires and capabilities of people to make non-home-based trips. Where they make them is, of course, another question. The authors would estimate zonal productions of non-home-based trips in two steps, first by determining area-wide totals, and second by allocation of these totals to production zones. Analysis of home interview sample data, again using each sample as an independent observation, would be employed to relate number of non-home-based trips made by household members (perhaps in more than one trip purpose category) to various socioeconomic and land-use parameters of the household. Solution of the resulting equation using area-wide forecast averages of the independent variables would produce an estimate of the number of non-home-based trips to be expected per household in the future. Multiplication by the forecast number of households would, of course, produce the total number of resident non-home-based trip productions in the area. Nonlinear variables could be handled in the same manner described for home-based trip production. Internal trips by external residents (which must be non-home-based) are usually not included in transportation studies. However, there is no reason why an estimate of the number of nonresident non-home-based trip productions in the area should not be added at this point.

An allocation function, to determine the number of trip productions within each zone in the area, could be derived from non-home-based trip production equations developed in some similar area where a full origin-destination survey had been made, yielding reasonably stable data on non-home-based trip production at the zonal level. Following this procedure, data collected in the particular area under study would be used to determine the overall significance of this type of travel, but the assumption would be made that the proportional influence of various parameters in explaining where non-home-based trips might be expected to originate is the same as in the similar area. This would seem to be a valid assumption in most cases.

An alternative means of deriving an allocation function would have to be employed, of course, if there were no reasonable non-home-based trip production equations up for adoption. One means of accomplishing this would be to break down the first step described—that of deriving an equation for the generation of non-home-based trips through relation to home parameters—into several origin purpose categories, and then distribute each purpose subtotal in accordance with the one zonal parameter that seems most reasonable.

The customary procedure in most transportation planning studies has been to derive an independent set of equations to relate trip attraction directly to various socioeconomic and land-use parameters. But O-D data from a small sample home interview survey will not provide a sufficient basis for this kind of trip attraction analysis. The authors recognize that this is a major disadvantage of the procedures suggested. Perhaps here again equations developed from a similar area, if such are available, could be used to advantage. Total attractions might then first be determined by recognizing that, over an entire area, total attractions must equal total productions plus or minus the net effect of trips crossing the external boundary of the area, and then be allocated to zones in one of the ways described for non-home-based trip productions. As far as non-home-based trips are concerned, it may not be unreasonable in some areas to assume that attractions equal productions in each zone. But, of course, such an assumption would not be valid for home-based trips.

The authors do not know of any actual transportation study where trip generation equations have been developed from unaggregated data at a household level and then used in the forecasting process. However, in a research project sponsored by the U. S. Bureau of Public Roads (2), multiple regression techniques were used to attempt to derive meaningful household trip production equations using data from 824 households spread all over the United States. Although the BPR-sponsored study was founded upon a limited nationwide sample, the findings demonstrated that this type of analysis had promise and ought to be pursued.

TRAFFIC ZONES

Traffic zones have traditionally been used as the basic unit of analysis for the entire transportation planning process. Trip data are usually coded to traffic zones, and socioeconomic and land-use data are usually inventoried on a traffic zone basis. Trip generation and distribution are also performed on a zonal basis. In traffic assignment, the trips to and from each zone are applied at a single point representing the centroid of trip end locations in the zone.

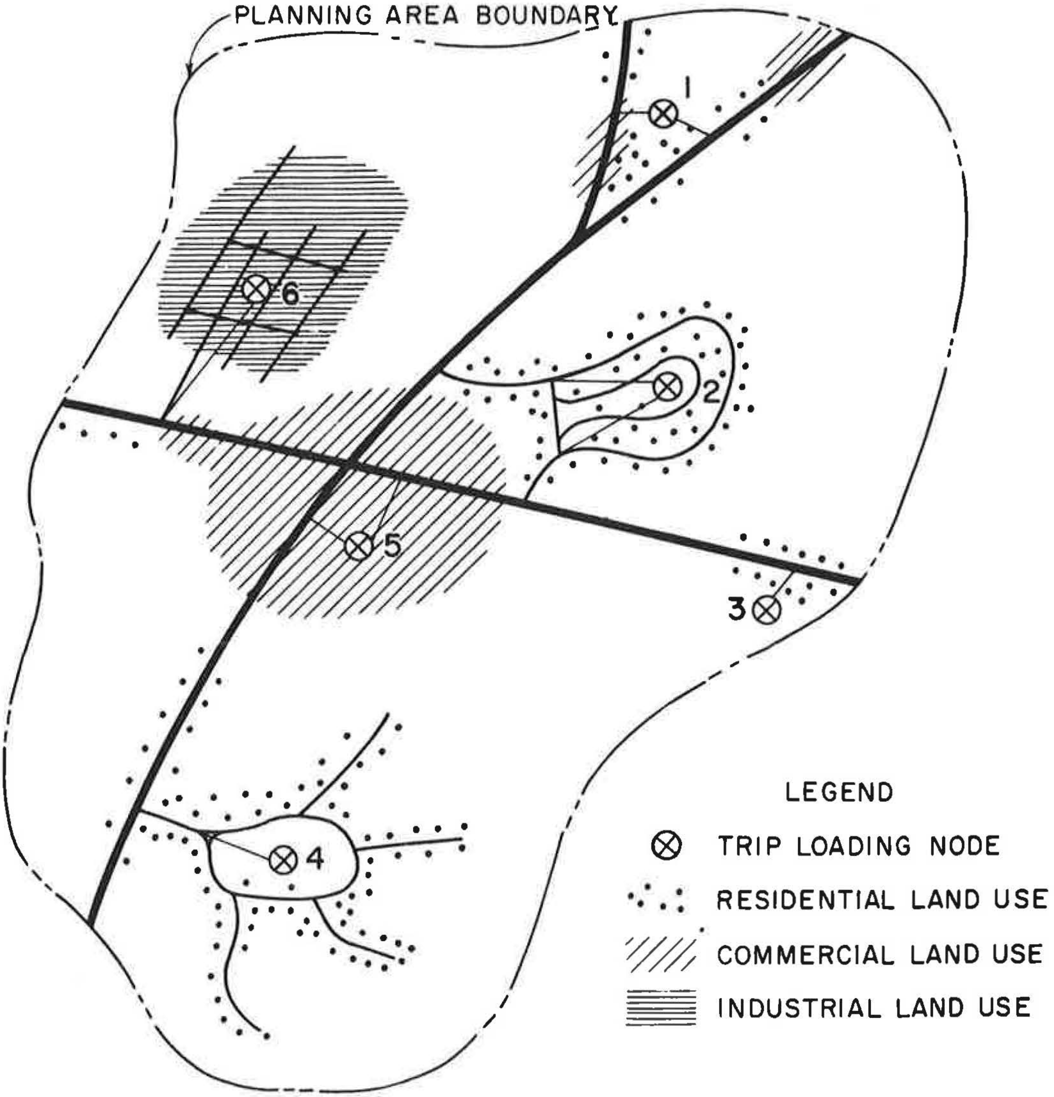
These different reasons for having traffic zones all translate into different criteria for delineating them, depending upon who is doing the traffic zone layout. Demographers are interested in availability of current and past population data; they prefer to have zone boundaries conform with areas already established by census authorities. The O-D survey data coders are interested in the ease with which addresses can be converted to traffic zone codes; they want to avoid running zone boundaries down the middle of streets, or they might even try to put entire streets in one zone or the other. The statisticians do not want the zones to be so small that sample data aggregated to the zonal level are statistically unstable. The transportation planners do not want the zones to be too large and are very concerned about the orientation of zones with respect to the transportation network. The trip generation analysts would be most concerned about homogeneity of zones, both from the point of view of character and of size.

Delineation of traffic zones in most transportation studies has been in accordance with criteria that did not permit wholesale adoption of areal units used for other purposes—for example, by the Census or by local planning bodies. Usually, when this occurs much of the value of past planning is lost and planning must be redone on the basis of the new system of traffic zones, which may not be logically and reasonably delineated from the point of view of regional planning. No layout of zones can ever satisfy everyone, but if traffic zones were not required to perform so many different functions at once, they should be able to perform certain functions better and more economically.

Traffic zones need not play a part in the trip generation analysis, and perhaps not even in the distribution model calibration process. If in neither, there would be no need to code O-D survey trip data to zones since we would be interested only in household characteristics, numbers of trips of various types (purpose, mode and time) made by household members, and the lengths of such trips—not specifically where the households were or where the trips went. Current trip generation and distribution would be simulated directly from household data. It would not be until this point then that some manageable unit of area larger than a household would be required. It is recommended that such a larger area, called a "planning area," be used.

PLANNING AREAS

Planning areas could be larger than traffic zones, and would be laid out primarily to serve the purposes of demographers, economists, geographers and planners. The transportation analysts would be concerned only with the degree of homogeneity of character for each of those parameters that affect generation of several trip categories being considered. Heavy industry and low-income, high-density residential development



TRIP CATEGORY	PERCENTAGE BREAKDOWN OF GENERATED TRIP ENDS BY LOADING NODE						
	1	2	3	4	5	6	TOTAL
Home Based Work							
Production	15	40	7	35	–	3	100
Attraction	5	2	1	2	20	70	100
Home Based Shopping							
Production	15	40	7	35	–	3	100
Attraction	10	–	–	–	85	5	100

Figure 1. Use of planning areas in travel forecasting.

could occur in the same planning area, but combinations of heavy industry and light industry, or of low-income, high-density residential development and high-income, low-density residential development within the same planning area would be avoided.

How then would the travel forecasting process be carried out? First, trip generation equations developed primarily from household data would be used to forecast the numbers of trip ends in each planning area. The forecast trip ends would then be disaggregated to trip loading nodes on some reasonable basis for each trip purpose category, as illustrated in Figure 1. The allocation procedure used to do this could vary widely in sophistication, just as the procedures used to allocate area-wide population forecasts, for example, vary widely in sophistication. But the objective in either case is the same—to get a fairly reliable forecast quantity for a larger area allocated to smaller areas.

Use of this concept would usually result in significant economies since all phases of the gathering of current data and the forecasting of future data would not need to conform to a single rigid system of traffic zones. Furthermore, the areal units used in the compilation and analysis of one class of data (e.g., population data) would not necessarily have to conform to the units used in the compilation and analysis of another class of data (e.g., employment data). Nor would the areal units used for current data necessarily have to conform to units used for forecasts. Several different breakdowns could be used for the forecast year, reflecting varying land-use development patterns. Of course, uniformity would be desirable where readily achievable, but the greater flexibility made possible by this approach may often provide advantages of economy and logic that far outweigh the relatively minor bookkeeping difficulties occasioned by having to keep track of more than one different system for subdividing the study area into workable planning area units.

Following the allocation of generated trip ends to loading nodes, trip distribution and traffic assignment would proceed in the normal manner. How many trip loading nodes are selected, where they are located and how they are connected to the transportation network would be left to the transportation analysts.

One of the important criteria in the selection of planning areas is the availability of data. Thus, for example, since census units are important to any study of population characteristics, they should be used as planning areas, at least for home-based trip analyses, whenever possible. Another criterion in the selection of planning areas is size. The size of individual planning areas must not be so great that a reasonably reliable allocation of any total pertaining to some characteristic of the planning area to various sectors within it cannot be made by inspection, on the basis of familiarity with the locale. It is very difficult to state this criterion in quantitative terms; its applicability in any given situation revolves around what is "reasonably reliable." This includes an appreciation of what impact a given level of imprecision in this allocation might have on the output of the transportation planning process, and also an appreciation of how significant a given level of imprecision in this allocation really is in comparison with the reliability of overall forecasts of regional activity.

DATA COLLECTION

The major advantage of the method of analysis proposed heretofore in this paper is that much less origin-destination survey information than is usually considered necessary will support it. Since data collection is one of the major items of cost in any urban transportation planning study, the implications of this fact alone are of great significance. But there are other possible opportunities for cost savings in the data collection phase that may be within reach also.

Home Interview Survey

Great care must be taken in the ordinary home interview survey to insure that samples are randomly selected from a complete statistical universe, and that sample data are expanded to account for the whole, even where there are gaps due to refusals and similar circumstances. These measures are essential when the purpose of the home interview survey is to observe and quantify a statistically sufficient sample of current travel so that the expanded sample represents the universe of current travel.

Using the type of analysis suggested in this paper, the purpose of the home interview survey would be to gather enough data to allow development of trip generation equations at the household level and to calibrate a trip distribution model.

Satisfactory calibration of a distribution model, such as the gravity model, from limited survey data appears to require that the small sample be uniformly distributed in a random fashion over the entire geographical area under study. The possibility of using a clustered sample of the same size was studied in a research project using data from the Pittsburgh Area Transportation Study (3). It was found that travel time factors developed from clustered survey data varied considerably from those developed from total study area data.

For the purposes of supporting the type of trip generation analysis proposed herein, the sample must provide sufficient observations within each socioeconomic grouping, but uniform sampling of all groupings and geographic dispersion of the sample are not required. As long as a small random uniform sample is required for gravity model calibration, it can be used, and probably will be sufficient, for sampling of household trip generation and related household characteristics as well.

If home interview survey data are to be used at the household level for the trip generation analysis, the question arises as to why the sample need be expanded at all. Current practice is to obtain trip length frequency from expanded survey data, but it would probably be just as reasonable to use trip data at the household level for this purpose also. Considerable savings could be effected if the sample did not require expansion. Moreover, if expansion is not required, do we really need the extensive and expensive measures traditionally taken in home interview surveys to insure complete, uniform, unbiased sampling of all segments of the universe? Since we are not seeking through sampling means a measure of the number of people in each socioeconomic group or in each geographic area, or a measure of area-wide trip generation or O-D trip patterns, must we encompass all corners of the statistical universe? If we are using utility records for sampling households, for instance, must we worry about the small percentage of households not served by the utility? Ordinarily we probably would because the proportion, though small, is indeterminate and we would have no basis on which to adjust expanded survey data.

These and other questions should be explored in the planning stage of the home interview survey. Perhaps collection of some additional data (e.g., stage in the family life cycle, availability of alternative transportation modes at times when specific trips were made, etc.) may be desirable to support the analysis of trip generation characteristics at the household level.

Alternative means for actually collecting the data should be investigated also. For example, several studies have reported success in making interviews by telephone (4). The Ohio Department of Highways uses a booklet which is dropped off at the sample address with a personal explanation of what is required, and later picked up and reviewed with the respondent to insure complete and accurate information.

Truck Survey

Heretofore not much has been said about analysis and forecasting of truck travel. The usual approach is to obtain travel data from interviews in a sampling survey in which the vehicle itself is the sample and an attempt is made to determine information concerning a particular day's travel.

There is a possibility that research will point the way toward a better understanding of truck trip generation leading to means of reducing the quantity of data required from an origin-destination survey. The initial step in this direction might be to separate trucks into three groups and deal with each separately:

1. First, trucks that are owned and used by individuals or families as personal vehicles should not really be considered as trucks at all in the trip generation analysis.

2. Second, a separate analysis should be made of travel by trucks that are primarily oriented toward providing services to households. These would include delivery trucks, repair trucks, refuse disposal trucks, mail trucks, and so forth. It should be possible to relate generation of trips by such vehicles to household characteristics; perhaps actual interviews would not be required at all for such vehicles, and sufficient data could be gathered to support a largely synthetic trip generation analysis through an expanded home interview survey or by special cordon counts around residential areas.

3. Third, all of the remaining trucks involved in the area's basic industries and businesses would undoubtedly have to be surveyed separately, and travel data analyzed and forecast on a zonal basis in the usual manner. It may be that there would be too much variation and too few observations from this group of trucks to support any sort of regression analysis, and that trip rate analysis by industry category would prove more satisfactory.

The authors' firm has recently completed origin-destination surveys in two urban areas in West Virginia (5, 6) in which part of the truck survey was conducted by mailed questionnaire. In both areas, questionnaires were mailed to owners of 100 percent of the non-fleet trucks (trucks registered to an owner who had no more than three trucks registered in his name). Fleet trucks were sampled and interviews conducted in the usual manner. In the larger of the two areas, a 38 percent return was received without making any follow-up mailing or telephone calls; a 27 percent return was obtained in the other area. The cost of conducting the mailed questionnaire survey was very much less than would have been required to select samples and make interviews in the normal manner. The average number of trips per interview was lower and the proportion of trucks making no trips on the travel date was higher for the mailed questionnaire survey than it was for the interview survey in both areas, but this is probably characteristic of the difference between non-fleet and fleet trucks. Research is clearly indicated to determine whether an uncontrolled sample of trucks, such as is obtained from voluntary return of mailed questionnaires, will yield unbiased trip generation and trip length frequency data.

An unusual procedure was employed by the authors' firm to collect truck travel data in a small urban area in New Hampshire (7). Here the home interview sampling rate was so great (1:5) as a result of the area's small size, that statistical stability of data relating to truck travel could be assured by obtaining origin-destination data for internal truck trips in conjunction with the home interview survey. This was accomplished with no particular problems, and at little increase in cost to the home interview survey. Special adjustments were made to account for wholly internal truck travel by external residents.

Roadside Interview Survey

In most urban transportation studies the collection of travel data at roadside interview stations on the cordon line surrounding the area represents a major portion of the total data collection effort. Careful design of questions to ask and forms on which to record answers is particularly important for the roadside interview survey because of the limited amount of time available for each interview.

An example of the kind of poorly worded question that should be avoided is the following, which oddly enough has become standard in many areas:

Question: Where is the car normally garaged?
Answer (circle one):

1. At origin inside cordon
2. At destination outside cordon
3. At neither
4. At origin outside cordon
5. At destination inside cordon

There are at least three things wrong with these answers: (a) they are confusing; (b) the specification of whether the origin or destination is inside or outside the cordon is redundant information; and (c) since the real purpose of asking the question is to find out whether or not the respondent is a resident of the internal area, if Answer 3 is given, do we know? Why not provide two answers—inside or outside—and let it go at that?

With regard to interview form design, the person who is organizing and preparing for the roadside interview survey should really have made some interviews himself, preferably in cold, wet weather at night with a large volume of traffic delayed. Many forms look fine in the office but are hard to use in the field. Poor form design is expensive, both in extra time spent in the field and in inaccurate, incomplete, or unreadable data brought back to the office.

A more rigid sampling technique might make it possible to reduce drastically the number of roadside interviews required. Roadside interview crews are customarily instructed to get all the interviews they can. There really is no control over the sample this way and we console ourselves that we make up for the resulting loss of statistical reliability by interviewing such a large percentage of the passing traffic that it can make no real difference. Fewer, more carefully selected samples would certainly result in lower costs for data collection and subsequent processing, and should produce just as reliable data.

Even greater savings would come from reducing the number of stations operated and the number of hours of operation, or from interviewing traffic in one direction only. To evaluate the feasibility of such measures, it is necessary to give careful consideration to the manner in which roadside interview data will be used in subsequent analyses and how external trips will be forecast.

Three types of trips are intercepted at cordon line stations—through trips, non-through trips by residents, and non-through trips by nonresidents. Non-through trips by residents are sampled in the internal home interview and truck-taxi surveys. Through trips and non-through trips by nonresidents can only be sampled at the cordon line.

The usual procedure for handling the duplication of resident trip data is to eliminate data pertaining to non-through trips by residents from the internal surveys. Non-through trips—by residents and nonresidents—are then distributed by the Fratar method or, in accordance with the latest recommendations of the U. S. Bureau of Public Roads (8), by a single-purpose gravity model. Through trips are customarily treated separately and distributed using the Fratar method. Duplication in the collection of data pertaining to through trips—each such trip has a chance of being intercepted twice, once at each point where it crosses the cordon line—is normally resolved by retaining all data collected but applying a one-half factor to them.

To obtain information about through trips it is necessary to interview on all the routes crossing the study area cordon line that carry an appreciable amount of through traffic. However, since through trips by definition cross the cordon line twice, we would be sure to obtain complete information if we were to interview traffic on such routes in one direction only—either inbound or outbound.

Trip data pertaining to resident travel are obtained in the internal interview surveys as well as the cordon line survey. The fact that we have duplicate sets of data to choose from is really somewhat of a luxury. If we did not have data from an external survey, we would certainly use what we had from the internal survey, probably without any qualms as to its adequacy.

This leaves non-through trips by nonresidents. What do we need to know about such trips to support the forecasting process that is normally used? We need to know how many trips there are. Insofar as total cordon line crossings are concerned, we would know this by interviewing only inbound or only outbound traffic, since total average daily inbound crossings must equal total average daily outbound crossings. In most cases it would be safe to assume that such an equality would hold for all stations individually as well as collectively.

Do we need to know trip purposes? If we follow the BPR recommendation and use a single-purpose distribution model, we do not need to know trip purposes. We must obtain from cordon line interviews information about trip lengths so that a trip length frequency curve can be developed for calibration of the distribution model. It would seem reasonable to assume that the trip length frequency distribution for inbound trips would equal the trip length frequency distribution for outbound trips. There are only two possible explanations why it would not: (a) triangular journeys with the inbound leg through one station, the outbound trip through another station, and the third leg outside the cordon line (it seems safe to assume that such journeys would not usually represent a significant part of the universe of non-through travel), and (b) triangular journeys with the third leg inside the cordon line (we customarily ignore such internal travel by external residents anyway).

It would appear then that, even without changing normal procedures for forecasting external travel, careful study should be given to the possibility of interviewing traffic in one direction only; probably inbound would be best. Through trips intercepted would not require a one-half factor; they would still be treated separately and distributed by the Fratar method. Non-through resident trips would still be deleted from the internal surveys and the assumption would be made that for every inbound non-through trip sampled at the cordon line, there is a matching outbound trip; non-through trips by residents as well as nonresidents would still be treated separately and distributed using a gravity model.

Further research is also required in the area of external travel forecasting. It is true that trip length frequency characteristics of external non-through trips may differ from those exhibited by internal trips, consequently requiring that external non-through trips be distributed separately. However, separate distribution requires separate sets of forecast trip ends, and the procedure used to split a forecast of total internal trip ends into those that must be distributed to other internal points and those that must be distributed to external points may be based on such questionable logic as to negate the benefits of separate distribution.

Other Travel Data Surveys

One disadvantage of the procedures suggested in this paper is that many of the opportunities available in the traditional approach for checking the completeness of data collection and the adequacy of models to reproduce current travel are lost. Thus, with current trip generation and distribution primarily simulated, and not enough origin-destination survey data to allow expansion to a universe of current travel against which the simulation models can be tested, how can we be sure that we have valid forecasting tools?

Comparison of ground counts with results of an assignment of simulated current travel will take on added importance as a test. But in designing an urban transportation planning study to incorporate some of the cost-saving techniques described in this paper, it will probably be necessary to conduct other data collection surveys of limited scope in order to provide the means for testing and evaluating the tools of travel forecasting in other ways as well.

Roadside interviews on a screenline running through the area would be particularly valuable in this regard. In some instances roadside interviews in major travel corridors, without necessarily forming a screenline, might be helpful. Collection of travel data at some of the principal employment and shopping centers or other major trip attractors in the area, perhaps by utilizing a postcard survey, should also be considered. A similar limited survey of transit riders might also be indicated.

CONCLUSION

The authors do not pretend that the cost-saving techniques discussed in this paper are any more than ideas, as yet largely untried. But many are felt to be worthy of immediate consideration, in planning for new transportation studies particularly. Other ideas presented are admittedly pure speculation and require careful evaluation through detailed research.

As stated early in this paper, the authors' main objective has been to stimulate the thinking of transportation planners in the direction of greater economy and to plant the idea that more economical techniques need not mean sacrificing any degree of reason and reliability in the planning process. Indeed the very techniques that will save money may lead the way to better planning.

REFERENCES

1. Ben, C., Bouchard, R. J., and Sweet, C. E., Jr. An Evaluation of Simplified Procedures for Determining Travel Patterns in a Small Urban Area. Highway Research Record 88, pp. 137-170, 1965.
2. Fleet, C., Stowers, J., and Swerdloff, C. Housing Trip Production: Results of a Nationwide Survey. Urban Planning Division, U. S. Bureau of Public Roads.
3. Heanue, Kevin E., Hamner, Lamelle B., and Hall, Rose M. Adequacy of Clustered Home Interview Sampling for Calibrating a Gravity Model Trip Distribution Formula. Highway Research Record 88, pp. 116-136, 1965.
4. Clark, James E., III, and Morin, Donald A. Securing Travel Data by Telephone Interviews. Traffic Engineering, Vol. 36, No. 9, pp. 47-49, June 1966.
5. Tippetts-Abbett-McCarthy-Stratton. Monongalia County Transportation Study, Morgantown, West Virginia (population 56,000).
6. Tippetts-Abbett-McCarthy-Stratton. Belmont-Marshall Counties Transportation Study, Wheeling, West Virginia (West Virginia portion only, population 106,000).
7. Tippetts-Abbett-McCarthy-Stratton. Dover-Somersworth Transportation Study, Dover and Somersworth, New Hampshire (population 30,000).
8. Calibrating and Testing a Gravity Model for Any Size Urban Area. Office of Planning, U. S. Bureau of Public Roads, Oct. 1965.

Discussion

JAMES J. McDONNELL, Chief, Urban Transportation Branch, U.S. Bureau of Public Roads—As in any other endeavor, there is a need in the urban transportation planning profession to continually evaluate the procedure being used in terms of both technical adequacy and cost. Ideally, this evaluation process should be made as standard operating procedure.

The ideas suggested by the authors represent a very commendable start in the evaluation process. They could conceivably form the basis of a research program taking into consideration, of course, the previous research conducted. Some of this research is identified in the paper.

The authors' main objective was to stimulate the thinking of transportation planners in the usefulness of the procedures discussed. In my case, they have succeeded in this objective. As I see it, the procedures identified for possible change are oriented around cost savings in the trip and socioeconomic portions of the data collection phase, and the trip generation and distribution portions of the analysis phase. In order to make a full evaluation, it would seem appropriate to also include traffic assignment procedures. Furthermore, speaking from the technician's point of view, it would seem more desirable to evaluate the procedures primarily on technical consideration and relegate cost to a secondary consideration. It is my feeling that the dollars that could be saved with more streamlined procedures would be greater than those saved by optimizing procedures that we all agree are not as perfect as we would like them to be.

The element of work that would be most affected by a reduced sample of home interviews or a sample that does not measure the universe of trips would be the analysis of nonresidential trip ends. When dwelling unit interviews are analyzed on a one-by-one basis, it is possible, as the authors have pointed out, to determine trips per dwelling unit, trips per car, trips per person, and the other commonly used residential trip generation rate factors; however, when limited surveys are used for a nonresidential trip generation analysis, it has been found that there is difficulty in establishing a stable universe for nonresidential trip ends. Therefore, it would be necessary to determine, for example, shopping trip ends from some other survey oriented to that need. Work trip ends could be determined from interviews conducted at work sites, provided there is the necessary cooperation. I contend that such an approach would be unsatisfactory to the statistician and also very expensive.

As the authors point out, the home interview survey has been with us for a long time. I look on its durableness as an indication of its merit. When an urban area is contemplating improvement programs that will result in the investment of millions of dollars of public funds, the dollars that go into a data base seem to me to be money well spent. It can be argued that the introduction of simulation models should have made our need for voluminous travel data less essential. Well, it has for certain areas of the process such as residential generation and trip distribution; however, the data also allow us to take a much more sophisticated look at other parts such as modal split and nonresidential trip generation analysis. These areas tax the data stability of even the "comprehensive" home interview survey data.

The procedures suggested by the authors show much promise when used in the continuing phase of the urban planning process. The continuing phase consists of five elements. They are surveillance, continuing reappraisal, service, research, and annual report.

The surveillance function is basic to the entire continuing process. Yearly maintenance of land-use and socioeconomic data as well as essential traffic and transportation data is the key to the continuing process. A trip end estimate based on changes in population and employment and other factors that have been determined in the initial phase to be significant in trip generation should be made on an annual basis. If this surveillance is done, then the procedures suggested by the authors could be developed to update and reevaluate the models and plans developed in the initial study.

Checks could be made of the models developed in initial studies by assigning the resultant trips to a current network and comparing them, as the authors suggest, to ground counts accumulated across screenlines and to vehicle-miles of travel checks by district and by facility type. At that time, the surveys required should be oriented to the refinement of initially developed models. It may be decided that the trip distribution model is the model that needs refinement and not the trip end models. If such is the case, then a survey could be developed to satisfy the need and for the reevaluation work, the old trip end models could be utilized. Such models may require adjustment in subsequent years and at that time a survey to satisfy a single purpose could be developed.

Specifically then such surveys would be structured around the solving of problems that have been identified in the reappraisal element of the continuing process. Any such methods should require a careful evaluation by the entire staff involved in the urban planning study.

These authors have brought thoughts and ideas to this forum. They should be evaluated in detail by other researchers in this country and the results of such work should be presented at future meetings of the HRB. It is through such research that worthwhile planning methods will develop for the use of persons conducting urban transportation planning studies.

RICHARD J. BOUCHARD, Director, Rhode Island Statewide Planning Program—Cost-saving techniques for O-D surveys have been a familiar topic of discussion at HRB meetings in recent years. In 1960, Robert Davidson, then with Boston University, reported on a very small sample survey which was used to develop a trip distribution model for the Boston region (9). Since that time, a large number of reports (for example, 10, 11, 12) have been presented documenting the validity of the smaller samples for purposes of developing such distribution models. But it appears that continuing discussion along these general lines is still necessary because even today large sums of money and, perhaps more important, large quantities of precious time are being spent in conducting, adjusting, and analyzing large-scale surveys.

The authors have set forth a number of suggestions which are said to reduce the time and costs involved in such surveys. While many of these proposals seem to be well presented and objectively discussed, the significance of some of them warrants appraisal.

The authors, for example, suggest that because of the rapid development of trip distribution models, current data needs are governed by the trip generation phase of the transportation planning process. They recommend that generation analyses be conducted on a household-level basis rather than on a zonal-level basis as a more realistic means of developing standardized trip generation equations which could be calibrated for any area of the country. While this recommendation may merit further exploration, its significance may well be questioned.

This discussant believes that trip generation has been made overly complicated because of a desire to "explain" too precisely the interrelationships between travel and the various characteristics which supposedly generate travel. Seven, eight, and even nine variable equations to generate total person work trips are not uncommon today. These equations have been justified because they explain an absurdly large percentage of the variation in those trips. On the other hand, two variable equations used to explain the same type of trip-making have also been used, but with more limited success in terms of their ability to explain current variations in trip-making.

The point, however, is that if the same statistically reliable measures are accepted in the trip generation phases as are already presumably accepted in the distribution and assignment phases, perhaps more rapid progress could be made in developing trip generation models which are standardized, at least to a degree comparable to present trip distribution and assignment models.

Three factors seem to support a less rigid statistical analysis of trip generation equations. First, it is certainly debatable whether the increased number of independent variables required to enhance the statistics of the equations are justified when one keeps in mind that all of these independent variables must be forecast 20 or 25 years hence. Second, it is doubtful whether the basic survey data can justify the attainment of rigid statistical results. Third, it would appear to be easier and more valid to compare one or two variable equations from various study areas throughout the country than it would be to compare equations with a larger number of variables. And this comparison is necessary if a standardized trip generation theory is ever to be developed.

So while the suggestion that something must be done to reduce the data needs of the trip generation phase is valid, the key to this reduction may well lie in acceptance of a lesser amount of statistical reliability. Once this notion has been accepted, then secondary improvements, such as use of the household rather than the zonal level, would be worthy of investigation.

In line with this, another thought might also be registered. The authors have correctly suggested that there are certain areas where the recommended procedure fails—notably with the production and attraction of non-home-based trips and with the attraction of all other types of trips. To combat this significant failure, the authors suggest that relationships be borrowed from other study areas. If these relationships are to be borrowed it would seem that the "standardization" previously mentioned must have been reached, at least to a greater degree than apparently has been reached to date.

The authors make a second principal recommendation in calling for the elimination of traffic zones and the establishment of so-called planning areas with multiple loading nodes.

It is somewhat unclear just what the significance of this proposal may be, particularly when considered in conjunction with the previous recommendation concerning trip generation at the household level. The authors suggest using the household-level trip generation equations at the planning area level, and disaggregating the results to trip loading nodes on some rational basis. If the basic premise that a household-level equation can be applied to an area representing a large number of households is accepted, then it makes little difference whether the equation is applied to the traffic zone level as is now customary, or is applied to the planning area level and disaggregated to a traffic loading node as suggested by the authors.

The authors suggest that their procedure is an improvement because it would eliminate the customary concern about standard data collection units and it would be possible, for example, to assemble population data on a census tract level and employment data on the individual establishment level. This may be so, but the simple fact remains that trips must eventually be allocated to a loading node and this loading node must represent a small traffic drainage area if the traffic distribution and assignment processes are to yield realistic results. Consequently, it makes little difference whether the trips are disaggregated from the planning area level to the loading node as suggested by the authors, or the social, economic, and land-use data are disaggregated from the data collection area as is now customary where the collection units are not standardized.

The authors further suggest that this procedure is an improvement because the planning area boundaries could be changed over the course of time and therefore could be adjusted to better reflect alternate land development patterns in the forecast year. The same course of action is possible with traffic zones today if the basic premise that the household-level equations can be applied to an areal unit representing large groups of households is accepted. In other words, whether you deal with a planning area and multiple loading node concept or with a traffic zone concept appears to be somewhat immaterial.

The authors make several other suggestions in their paper which bear some comment. They suggest, for example, that it may not be necessary to expand survey data to the total universe. This suggestion appears valid. In fact, in two recent surveys (13, 14) conducted in Rhode Island, the data were not expanded and the results have been entirely favorable. Connecticut has also followed the same procedure in at least two surveys (15, 16), and similar conclusions were reached.

A telephone survey and postcard questionnaire are mentioned by the authors as possible alternative methods of collecting data. Experience in Rhode Island indicates that both methods are satisfactory for collecting origin-destination survey data (17) and that significant cost savings can be realized by employing such techniques.

The authors' suggestions on truck surveys and roadside surveys, for the most part, seem appropriate and worthy of implementation without additional investigation.

Finally, the authors make the point that, as sample sizes are reduced, the value of a good volume-counting program becomes more critical. The importance of this statement, regardless of the amount of O-D data collected, must be recognized by any planning program which desires to be active in providing highway design figures.

In summary, the authors have presented a number of suggestions, many of which are quite valid and quite significant. These should be implemented with a minimum of further delay. They have also presented two suggestions which may be questionable as to their significance and validity. These should be further investigated in the near future, perhaps by the authors, and the resulting facts and figures presented as soon as possible.

References

9. Davidson, Robert G. Developing a Traffic Model With a Small Sample. HRB Bull. 297, pp. 106-108, 1961.
10. Kudlick, Walter, and Fisher, Ewen S. Use of Pre-Interview Trip Cards in Developing a Traffic Model for the Hamilton Area Transportation Study. HRB Bull. 347, pp. 186-202, 1962.
11. Smith, Bob L. Gravity Model Theory Applied to a Small City Using a Small Sample of Origin-Destination Data. Highway Research Record 88, pp. 85-115, 1965.
12. Wiant, Rex H. A Simplified Method for Forecasting Urban Traffic. HRB Bull. 297, pp. 128-145, 1961.
13. Manual of Procedure for Telephone Home Interview Recreation Surveys. Rhode Island Statewide Planning Program, Providence.
14. Manual of Procedure for Post Card Transit Ridership Surveys. Rhode Island Statewide Planning Program, Providence.

15. Barnes, Charles F., Jr. Integrating Land Use and Traffic Forecasting. HRB Bull. 297, pp. 1-13, 1961.
16. Southeastern Connecticut Area Transportation Plan. Connecticut State Highway Department, Wethersfield.
17. Preliminary Analysis of 1966 Post Card Transit Survey. Rhode Island Statewide Planning Program, Providence.
18. Crevo, Charles C. Synthetic Trip Generation and Distribution—A Practical Development and Application. Rhode Island Statewide Planning Program, Providence.
19. Preliminary Analysis of 1965 Telephone Recreation Survey. Rhode Island Statewide Planning Program, Providence.

AUSTIN E. BRANT, JR., and DANA E. LOW, Closure—The authors wish to express their appreciation to the discussants for their thoughtful consideration of the paper.

Mr. McDonnell emphasizes the important point that any changes in procedures used in the transportation planning process must be evaluated from the technical standpoint as well as on the basis of costs. The aim should be to improve quality while reducing costs, but we cannot ignore the fact that highly sophisticated procedures which utilize large volumes of data and complex methods of analysis may not really result in a better product.

Mr. McDonnell suggests that special surveys would be needed to establish a stable universe for nonresidential trip ends. The data already available should, of course, be fully utilized. Area-wide totals for attracted trips can be determined since total attractions must equal total productions, corrected for trips crossing the area boundary. Detailed employment data are available in the records of employment security agencies. Research projects, such as Keefer's studies of airports, shopping centers and industrial plants (20), can be used to estimate attractions. While an approach which uses data from many different sources and which accepts models developed in other studies may not be statistically satisfying, it should be more than adequate for engineers and planners who realize that the entire transportation planning process is based on forecasts of socioeconomic data 20 to 30 years in the future. The degree of error inherent in such forecasts will far outweigh any statistical errors introduced by using streamlined procedures.

The authors have not attempted to indicate that the home interview survey has been in use so long that it is now outmoded. Dwelling unit interviews form an essential part of the procedures we have proposed. The use of models, however, reduces the requirements for the type of data obtained in the home interview survey. Admittedly, the proposed procedures may reduce the statistical validity of simulation for certain types of travel, but very often these types form only a small part of total travel and have a very limited effect on forecast design hour volumes.

Mr. McDonnell's suggestions that the proposed simplified procedures be applied to the continuing phases are most appropriate.

Mr. Bouchard points out that trip generation equations can be made overly complicated by incorporating an excessive amount of variables. This is very easy to do, of course, in multiple regression analyses. After working with computers for some time, one may lose sight of the common-sense relationships between cause and effect. The development of objective standards for trip generation would seem to be a fertile field for research. Mr. Bouchard also brings out the point that all independent variables must be forecast far into the future. These forecasts may range from the hopes of land-use planners for the types of development that they would like to see occur in the future to trend-line extrapolations of what has occurred in the past.

Mr. Bouchard comments on the proposed use of planning areas in lieu of traffic zones. When small traffic zones are used, those responsible for land-use planning are required to subdivide their data collection and forecasts into small geographic areas.

While land-use planners can forecast the type of developments which will occur in large areas, it is unreasonable to expect them to select exact sites for industrial developments, shopping centers and the like. Forcing land-use planners to use traffic zones does not improve the accuracy of the result and may be misleading. What is proposed is that land-use planners forecast for areas which are within their capabilities and that subdivision into smaller areas, or allocation to loading nodes, be done as part of the transportation analysis. The authors believe that it does make a difference in cost whether or not socioeconomic data are compiled and forecast on the basis of small areas, defined in a consistent manner for all variables. The disaggregation can be accomplished at less cost as part of the transportation analysis. Perhaps the difference is one of semantics, but the proper use of words can often clarify concepts.

Mr. Bouchard reinforces the statements in the paper concerning the value of a good volume-counting program. The authors have not been engaged in any study where there were too many counts of traffic volumes or transit passengers.

The authors feel that they have accomplished their first objective—to stimulate the thinking of experienced transportation planners concerning the methods and procedures they use, even though these methods and procedures have been in use for many years.

Reference

20. Keefer, Louis E. Urban Travel Patterns for Airports, Shopping Centers, and Industrial Plants. NCHRP Report 24, 1966.