TRANSPORTATION RESEARCH

RECORD

No. 1477

Planning and Administration

Statewide Travel Surveys, Traffic Data Collection, and Urban Travel Patterns

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1995

Transportation Research Record 1477
ISSN 0361-1981
ISBN 0-309-06115-6
Price: $23.00

Subscriber Category
IA planning and administration

Printed in the United States of America

Sponsorship of Transportation Research Record 1477

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Chairman: Thomas F. Humphrey, Massachusetts Institute of Technology

Transportation Forecasting, Data, and Economics Section
Chairman: Mary Lynn Tischer, Virginia Department of Transportation

Committee on Transportation Data and Information Systems
Chairman: Arthur B. Sosslau, Comsis Corporation
Secretary: James J. McDonnell, Alexandria, Virginia
David Preston Albright, Marsha Dale Anderson, Jack A. Butler, Patrick R. Cain, Ed J. Christopher, Gary Q. Coe, Keith A. J. Crawford, Kenneth J. Dueker, Patricia S. Hu, Martha M. Johnson, Charles A. Lave, Alan E. Pisarski, Charles L. Purvis, Phillip A. Salopek, Eddie Shafie, Howard J. Simkowitz, Ronald W. Tweedie, Xu Weici, George F. Wiggers

Committee on Transportation Planning Applications
Chairman: Gordon A. Shunk, Texas Transportation Institute
Rick Donnelly, Jerry M. Faris, David T. Hartgen, Paul Hershkowitz, Charles E. Howard, Ron Jensen-Fisher, George T. Lathrop, Eugene J. Lessieu, Richard S. Marshment, George L. Reed, G. Scott Rutherford, Eddie Shafie, Arthur B. Sosslau, Montie G. Wade, W. T. Watterson, Edward Weiner, Debra Ann Whitmore, George V. Wickstrom, William A. Woodford

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
Chairman: Jerome W. Hall, University of New Mexico

Users and Vehicles Section
Chairman: Conrad L. Dudek, Texas A&M University

Committee on Traffic Records and Accident Analysis
Chairman: Paul B. Jovanis, University of California-Davis
Secretary: Jeffrey E. Paniati, Federal Highway Administration
Sherry S. Borener, David J. Bozak, Thomas E. Bryer, Myung-Soon Chang, Benjamin V. Chatfield, Ted Chira-Chavala, Barbara Hilger DeLucia, Mark Lee Edwards, C. Arthur Geurts, Alan F. Hoskin, Patricia S. Hu, John J. Lawson, Joseph C. Marsh IV, Judson S. Matthias, Ronald C. Pfefer, Carol Lederhaus Popkin, Howard S. Stein, John G. Viner, Phyllis E. Young

Transportation Research Board Staff
Robert E. Spicher, Director, Technical Activities
James A. Scott, Transportation Planner
Nancy A. Ackerman, Director, Reports and Editorial Services
Naomi Kassabian, Editor

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1994.

Transportation Research Record 1477

Contents

Foreword

Various aspects of data collection strategies for transportation planning are discussed in this volume. Wynter and Casavant et al. discuss data collection techniques for freight transport, including stated preference surveys and personal interviews. Techniques for data collection for input into the development of statewide travel demand forecasting models are covered by Virkud and Keyes and Crevo et al. Other topics include measurement of time-based traffic events using video recording, development of a Global Positioning System for traffic safety, work trip analysis using household survey techniques, application of the 1990 Census Transportation Planning Package, and estimation of pass-by trips using a license plate survey.

Stated Preference Survey for Calculating Values of Time of Road Freight Transport in France

LAURA M. WYNTER

A survey based on a stated preference experiment aimed at obtaining information about road freight transport subsequently useful in evaluating freight transporters' critical values of time is presented. The initial hypothesis in designing the survey was that values of time may vary significantly across the population; the experiment was thus designed so as to permit the evaluation of these critical values for each individual, as well as the identification of possible dependencies of the values on other trip attributes. The scope and methodology of the survey as well as an overview of the results are presented.

The first efforts in assigning values of time can be found in the economic literature in the 1960s (*1*). The formulation considered by Becker (*2*) in 1965 stated that the utility of an activity included both the elements obtained directly through the activity, the time taken, and possibly intermediate activities. The individual then tries to maximize his utility subject to budget and time constraints.

According to Gronau (*1*), applying this theory to transportation requires suppressing a common hypothesis of the more general application to economics. That is, it can no longer be assumed that the characteristics of the activity can take on a continuous range of values when the activity in question is a trip; indeed, the possible monetary costs and travel times are fixed by the infrastructure for a particular trip. Thus, to change the ratio of price to time for the trip in question, the user is required to choose another mode or itinerary. Gronau concludes that the different choices in transportation are thus perfect substitutes. This property will allow the evaluation of the user's *critical point*, where the critical point separates the efficient combinations of price and time combinations that an individual is willing to accept from those for which the trip is too costly (or too slow) to be undertaken.

The value-of-time (VOT) parameter, in general, transforms variables measured in time units to monetary units. Mathematically, it is simply a conversion factor that enables one to take a summation of both time and money constraints. Economically, however, it describes the value given by an individual for a unit time gain. This latter interpretation incorporates a number of hypotheses of human behavior. Two such hypotheses and the manner in which they are addressed are discussed next.

First, however, the use of a quantitive VOT is addressed. In general, the value obtained will represent the individual's valuation of time savings, rather than of absolute time. It is clear that the range of values considered is important in this case; a 2-min time savings is not necessarily equal to 1/5 of a 10-min savings, which may not be equal to 1/10 of a 20-min savings. Indeed, some time savings will be useful when they can be directly applied to another activity, and others will have no value. In the case of freight transport, however, where the necessity of the trip and cost-minimizing behavior are assumed, it may be assumed that most time savings lead to a savings in money (through more efficient use of trucks and drivers).

The two hypotheses concern the users. They are (a) the existence of a range of VOT across the population and (b) the existence of a correlation between the VOT and other trip attributes. Although a single VOT for the entire population is often assumed in travel demand models, some French transportation researchers have, since the 1970s, assumed that there is a continuity of VOT in passenger travel [see, notably, work by Abraham and Blauchet (*3*) and Marche (*4*)]. In this stated preference experiment, these two hypotheses for freight transport will be addressed by an appropriate survey design.

STATED PREFERENCE SURVEY

Developed in the context of market research, in which the demand for new products needed to be estimated, stated preference (SP) techniques have responded to the need of researchers to create and control a (hypothetical) situation and then to measure the responses. In the case of transportation, researchers can use these techniques most efficiently to evaluate the possible demand for potential new services (including price increases) and infrastructures.

SP techniques are known, however, to possess diverse biases, including justification or overestimation as well as a "political response" bias (*5*). The justification bias stems from a user's trying to "justify" a current choice; the user may thus be overly reluctant to indicate a change, even if he or she will later accept it. Similarly, an overestimation bias occurs when a user is initially too enthusiastic and accepts too readily a change that he or she would later refuse. Finally, a political response bias results when an individual responds to a question in such a way as to influence the outcome of the survey. (This may be the case in particular when the questions involve possible price increases, as they do in this experiment.) Nevertheless, in spite of the possible biases, SP experiments offer an important tool for the analysis of behavior.

An SP survey is the only feasible way of estimating the critical values of time. Revealed preference (RP) data describing users' present travel costs and trip times do not provide enough different ratios of (toll) price to travel time since, for a particular origin-destination pair, there is usually not more than one or two different (toll) price-time combinations; the critical ratio is impossible to estimate from only one or two points. An SP survey, however, allows the isolation the of the effects of price and travel time changes so as to obtain directly the user's critical VOT.

Institut National de Recherche sur les Transports et leur Sécurité, 2 avenue du Général Malleret-Joinville, 94114 Arcueil-Cédex, France.

Survey Design

SP surveys can be classified into three groups according to the types of questions posed: ratings, rankings, or discrete choices. The first type is given on a numerical scale on which the individual indicates the degree to which the option is preferred (e.g., 10 implies a strong preference, and 1 indicates that the option is unattractive). Rating techniques show the relative preference for the options with respect to the others presented by asking the individual to order them according to increasing attractiveness. Discrete choices present two options between which the user chooses.

The type of experiment chosen depends on the desired analysis method. Rating and ranking exercises give a larger quantity of information than discrete choices but are accompanied by less realistic results and thus higher random error. (For example, a ranking exercise naturally assumes equally spaced intervals between the options; this may hide the individual's true preference levels. Rating exercises, while overcoming this problem, may be too demanding when the behavior in question is difficult to quantify exactly.) In this study, the price or trip time at which a particular itinerary (toll highway or state road) is valued equally to the other is of interest; a single point is thus desired for each individual. Consequently, a discrete choice experiment was best adapted to the objectives of this study.

In an SP survey, correlations between the independent variables, often a problem in RP data, can be avoided by a careful selection of the questions presented, particularly by ensuring that the attributes vary independently (*6*). In this survey, it is assumed that the main attributes of the autoroutes and state roads are well understood by the subjects and that it was not necessary to enumerate them explicitly. The tolled highway ("autoroute") is considered a fast road, flat and with few perturbations, but accompanied by a high price. The state roads often have traffic lights in towns and other perturbations including steep ramps (in the mountainous regions) and sharp turns (these affecting significantly the speed and fuel efficiency of trucks), but are free and offer better opportunities for the driver to eat and rest.

Assuming that these characteristics were well understood, the survey questions were focused on the effects of modulating them. According to Pearman et al. (*5*), to be valid, the options presented in an SP survey should (a) appear plausible, (b) correspond to previous experience of the subject, (c) permit competitive trade-off decisions, and (d) cover the possible scope of responses. In this survey questions were presented concerning toll increases on the autoroutes and travel time increases on the state roads, as well as additional questions about gasoline price increases and toll decreases on the autoroutes (for those who take the free roads only). Aside from the last question, which mentions an option not genuinely plausible, the questions satisfy the first three guidelines cited above. The last guideline, that the values presented cover the scope of possible responses, may not hold; however, a more exhaustive questionnaire would not have been feasible in the context of a telephone interview. (It was decided to use telephone-based surveys because it was feared that the mail-back rate and time delay among professional transporters would be exceedingly low.)

Questionnaire

The first job of the survey taker, after introducing the project and verifying that the transporter did indeed carry out interurban road freight transport, was to ask the transporter to recall a typical recent (or regular) trip. The rest of the survey was divided into two parts: the first was a description of the interurban trip in question, including the type of merchandise, its packaging, the type of vehicle, approximate trip distance, origin and destination, itinerary, and time constraints on the trip. In addition, questions were posed as to the size of the transport company itself.

The second part of the survey contained the discrete choice questions aimed at evaluating the critical point between toll price and travel time for that transporter and that trip. This critical point could only be calculated when the transporter was forced to choose between a free state road itinerary and a faster, but high-priced, autoroute itinerary. If the choice of these two types of itineraries did not exist, the interview was not continued.

The discrete choice questions posed depended on whether the trip in question was made on the tolled autoroute or on the state roads. The survey taker provided the toll price of the trip on the autoroute and the travel time of the same trip on the state road. For those who currently chose the autoroute, the questions were devised so as to obtain the toll price at which they would no longer be willing to pay and would thus switch to the free state roads. On the basis of the price currently paid, three options could be proposed, starting with a choice between the state road and the autoroute at 125 percent of the current toll price. If that increase was accepted, the same choice was proposed with the autoroute toll at 150 percent, and last the toll price proposed was 200 percent of the currently paid toll if the subject had accepted the +50 percent increase. These levels were chosen to cover most of the expected trade-off points. An increase of less than 25 percent in the toll price would have been imperceptible for trips of medium distance, whereas an increase of over 100 percent may have been perceived as unrealistic. Though a larger number of levels would have served to more accurately identify the trade-off point, it was believed that the added fatigue and irritation on the part of the transporter would have negated these benefits by adding biases and random error.

For those who currently chose the free state roads, three sets of discrete choice questions were posed. The first presented increases in travel times on the free roads, and at each increase, the choice between that and the tolled autoroute was given. Four levels were proposed: a travel time of 110, 125, 150, and 200 percent of the current one on the free road. (A 10 percent trip time increase, unlike the analogous toll increase, was believed to be perceptible to the transporters.) The other two sets of questions were used as checks: three levels of toll price decreases and three levels of fuel cost increases were proposed. (It is known that trucks use significantly more fuel on state roads because of braking at traffic lights and on the ramps and sharp turns.) However, it was found that these checks were not very reliable; as suspected, the toll price decrease did not provide robust results (there was a much higher percentage of illogical responses), and many transporters did not appear to be concerned by a possible increase in fuel prices.

Sample Group

A random selection of interurban road freight carriers was obtained from a national directory. The subject of the interview was the fleet manager; in the case of very small companies, the fleet manager was often also the driver. Approximately 650 completed surveys were obtained, representing only 20 percent of the attempted telephone interviews. The low response rate can be attributed in large part to the competitive nature of the road freight transport industry in France. (In Paris, where competition is the most intense, the response rate was significantly lower than that in the smaller towns.) Nearly 40 percent of the 650 surveys needed to be eliminated; in some cases, both autoroute and state road itineraries were not available for the trip in question; in others, a significant policy response bias was evident. The number of usable questionnaires was thus reduced to a final sample group of 408.

Nevertheless, the remaining sample covered a wide geographic range: 58 of 95 French departments were represented as trip origins and 69 as destinations. Furthermore, the sample group characteristics were shown to well represent the target population. Table 1 shows a comparison of the sample with the true population of interurban road freight haulers. It can be seen that in terms of trip distances, the range of goods transported, and the current split between free state road and autoroute itineraries, the sample group provides a reasonable representation of the target population.

CRITICAL VALUES OF TIME

In the analysis of road pricing schemes, it is necessary to have a model based on the critical points at which the individuals being studied are indifferent to the options presented, that is, between the cheaper and slower (or less comfortable, convenient, etc.) option and the better but more expensive alternative. At a price inferior to this critical point, the individual will choose the better option, but once the price of the better option exceeds that individual's critical point, the slower, cheaper alternative will be preferred. The trade-off points obtained in the stated preference survey permit the calculation of each individual's critical VOT.

A shortest-path program over the French road network was implemented to provide the true distances, toll costs, and travel times of the itineraries obtained from the survey. In addition, vehicle operating costs were calculated on the basis of each transporter's reported vehicle type. Thus, for each individual, the generalized costs of the two alternatives at the trade-off point for the currently selected option could be set equal to one another and the critical VOT determined. Suppose that the generalized cost, G_{ij}, of transporter i on path j is given by

$$G_{ij} = c_{ij} + p_j + VOT_i\, t_j \tag{1}$$

where c_{ij} is transporter i's vehicle operating cost on path j, p_j the toll price, and t_j the travel time on path j, and setting

$$G_{iT} = G_{iF} \tag{2}$$

so that at the tradeoff toll price p_T for those who choose the toll road (or travel time t_F for those on the free roads), the two generalized costs are equal from the point of view of transporter i. Solving for each transporter's critical VOT_i,

$$VOT_i = \frac{(c_{iT} - c_{iF}) + p_T}{t_F - t_T} \tag{3}$$

The resulting values of time range from approximately 1 French franc/min to 36 francs/min. Looking ahead to the next section, it can be seen that the first hypothesis of a range of VOT across the population is strongly supported. This result is quite important since the majority of transport planning models use a single value for the entire population. Diversity of VOT will be discussed in more detail in the next section, as well as the second hypothesis, which concerned the correlation of VOT with trip characteristics.

ANALYSES OF INITIAL HYPOTHESES

If has been assumed from the outset that the value of travel time is more complex than its usual treatment indicates. An SP survey was devised to test various hypotheses, as discussed in the introduction of this paper. The first hypothesis was the existence of a range of VOT across the population of freight transporters. The second hypothesis concerned the variation of VOT with other parameters of the trip.

Diversity of VOT Across the Population

The first hypothesis was confirmed by the stated preference experiment. As stated in the previous section, a wide range of VOT was obtained for the sample population, from 1 to 36 francs/min at the high end of the spectrum. It was shown, furthermore, in the section on the stated preference survey, that the sample population represents well the desired scope of the study, that is, the interurban road freight transport industry.

The set of critical VOT can be visualized more easily by grouping them into a histogram, which gives an approximate density function (Figure 1).

In this histogram, each group has a width of 2 francs/min. It is easy to see that the VOTs vary more or less continuously over the population, and that, in fact, the pattern approaches a very elegant log-normal distribution. This empirically obtained result is particularly interesting, first because it contradicts the commonly held belief in a single (or a small number of) VOT in the population. Second, this result provides a distribution that is very easy to handle

TABLE 1 Analysis of Freight VOT Study Group

	Sample group	Target population (7)
% on toll road	80	76
Trip distances		
% trips 100-300 km	21	21
% trips 300-500 km	32	24
% trips 500+ km	38	25
Good types transported		
% Agricultural	36	30
% Petroleum	2	7
% Primary/semi-manufactured	13	12
% Manufactured	49	51

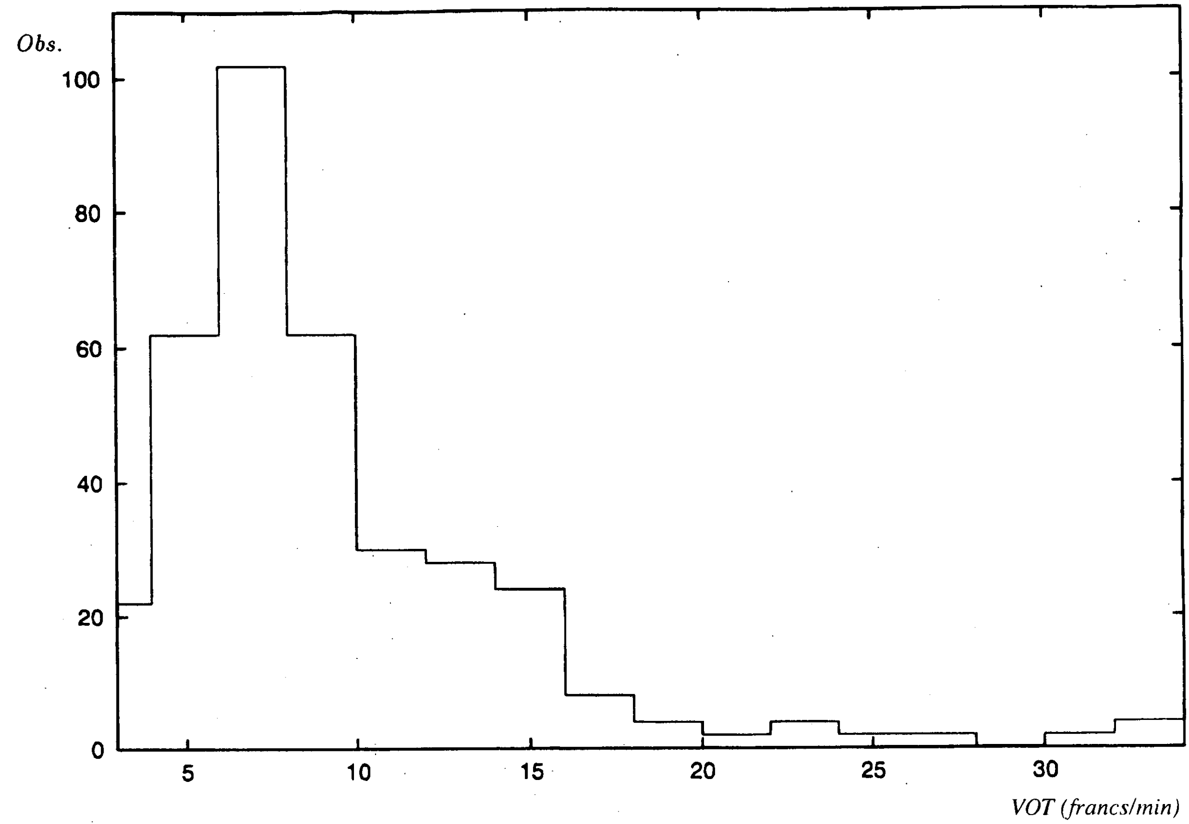


FIGURE 1 Histogram of critical VOT data points.

mathematically. In addition, it confirms a much-used hypothesis in France concerning critical VOTs. (Transport economists in France have believed for some time that passengers' VOTs follow a distribution similar to that of income in the population, which is a log-normal distribution.)

Such a distribution is furthermore relatively easy to incorporate into any mode choice or traffic assignment algorithm and thus can substantially improve the predictive power of the model. In the case of the logit model, Ben-Akiva et al. (*8*) showed how one can incorporate a continuous distribution of VOTs directly into the logit function. For the price-time traffic assignment model, Leurent discusses the use of such a continuous distribution (*9*), and an implementation is discussed by Morellet (*10*). For other models, mode choice or traffic assignment, VOTs can be chosen by a discretization of the VOT curve into VOT classes for use within standard multiclass models [see, for example, the work by Sheffi (*11*)].

Thus the optimal log-normal curve representing the data from this survey was calculated. The mean and standard deviation of the data set are given by

$$\mu_{VOT} = 8.65 \qquad \sigma_{VOT} = 5.94 \tag{4}$$

where these values are in French francs per minute. Estimating the parameters m and s of the log-normal function $\ln(v) \sim N(m, s)$,

$$F(v) = \frac{1}{v\sqrt{2\pi s}} \exp \frac{[\log(v) - \log(m)]^2}{2s} \tag{5}$$

where the resulting parameter values are

$$m = 7.1257 \qquad s = 0.2726 \tag{6}$$

The resulting distribution, normalized to 1, is shown in Figure 2.

The expected value of the log-normal distribution is given by

$$E[v] = \exp\left(m + \frac{s}{2}\right) \tag{7}$$

Thus, it has been shown clearly that the hypothesis of a single (or small number of) VOT in the population does not hold. Furthermore, the use of a larger number of VOTs permits an effective testing of road pricing scenarios, in that some individuals will be incited to switch to the lower-cost routes (or times of day) as the toll prices are increased, while those users having elevated VOTs will prefer to stay on the faster autoroutes. [See work by Leurent (*9*) for more details.]

Variation of VOT with Trip Parameters

Considering the second hypothesis, that is, that the VOTs may vary with trip characteristics, examination of the VOT over characteristics such as the type of merchandise transported, the time constraints of the trip, the vehicle type, the origin-destination pair, and the trip distance reveals that the only significant dependence in the data sample is on trip distance.

This result is quite important in that it permits the quantification of a phenomenon that is often observed. That is, for longer-distance trips, a larger portion of the population (in both passenger and freight transport) choose the high-priced toll roads over the free, but slower, state roads. The use of a distance dependency permits quantification of the degree of this effect by allowing the mean VOT to vary with trip distance. This result is clearly more significant in interurban travel than in urban transport modeling, where trip distances are contained in a tighter range.

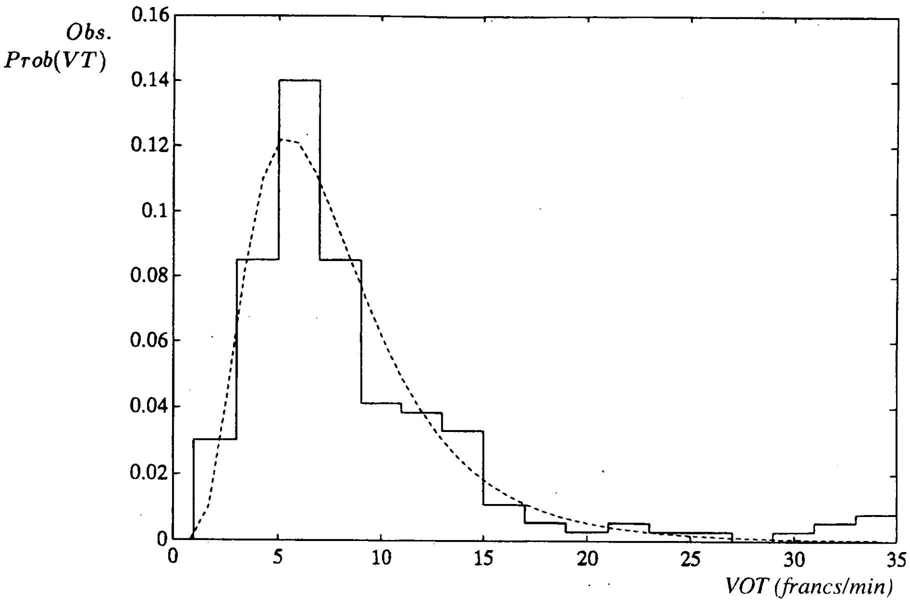


FIGURE 2 Optimal log-normal distribution of critical VOT.

The mean VOT in this case can be shown to increase linearly with distance. Estimating the line given by these points by an OLS method gives

$$\mu_v(\text{dist}) = 5.39 + 7.83 \frac{\text{dist}}{1000} \quad (8)$$

where the linear goodness-of-fit measure $r^2 = 0.75$.

Again, this result is directly usable. The mean VOT can be simply calculated for each origin-destination pair in a model by first obtaining the shortest path distance and substituting it into a formula such as Equation 8. Furthermore, this result can be combined with the results of Hypothesis 1, that is, the continuous distribution of VOT. In particular, a new log-normal distribution can be written in which the function depends not only on the VOT parameter but also on the trip distance. Figure 3 illustrates this extended VOT distribution.

Further details on the calculation of these VOT distributions can be found elsewhere (*12*).

CONCLUSIONS

A stated preference survey was designed that enabled calculation of freight transporters' individual, critical VOTs. That is, the trade-off points between increased trip cost for a faster route and cost savings at the expense of a longer travel time were obtained. Because the questions were posed for a particular trip, it was possible to obtain complementary information that enabled the examination of the dependence of VOT on other trip characteristics.

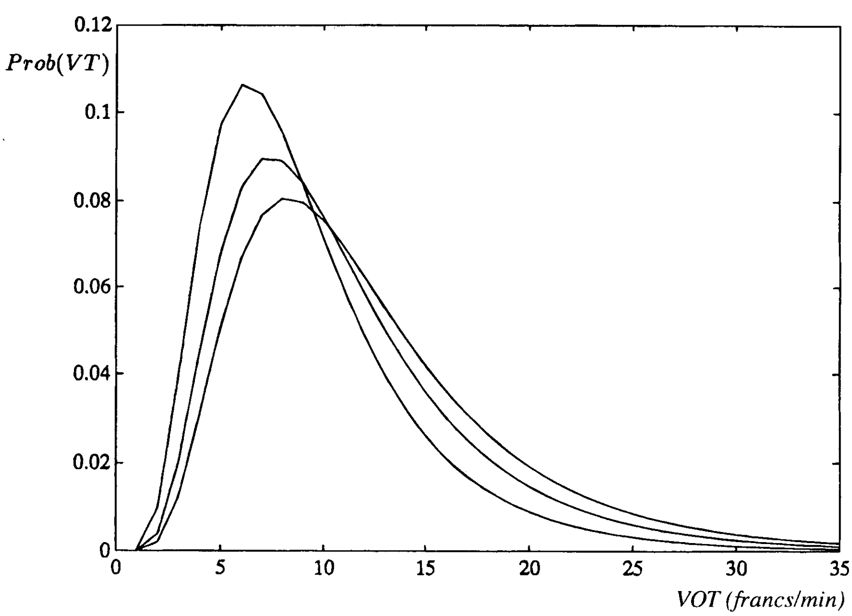


FIGURE 3 Distance-dependent log-normal distribution of critical VOT.

In doing so, it was possible to show that two commonly held beliefs about VOT are not, in fact, valid. For the case of interurban road freight transport, it was showed, first, that the VOTs of French transporters can be described by a log-normal distribution and thus that the use of a single, mean VOT is an oversimplification.

As such, transport models that use a single, mean VOT in their generalized cost functions must be called into question. This is particularly true when one of the objectives of the transport model is to analyze traveler response to changes in toll prices. Leurent (*9*) has shown that such a multi/valued VOT (continuous or discrete) is, in fact, necessary for effective testing of road pricing scenarios. Indeed, the use of a single, mean VOT in such a model means that all travelers react in an identical manner to a price increase. Clearly, it is not possible to determine the effect of toll changes on traffic flows without a VOT dispersion across the population, allowing a multiplicity of driver reactions along each path.

In addition, it was shown that these mean VOT are not independent of trip characteristics but do vary, particularly with trip distance. In this case, the VOTs of French road freight carriers were shown to vary linearly with trip distance. This confirms the general observation that a larger percentage of carriers are found on the tolled autoroutes when they are traveling long distances. Furthermore, it indicates that the use of a (marginal) VOT as a linear parameter of time should be valid.

Again, it can easily be seen how this factor would improve the predictive ability of traffic assignment models. When traffic is assigned to long-distance itineraries, a set of flows would be obtained (across tolled and untolled itineraries) that is different from the set that would be obtained for a medium-distance itinerary. More specifically, an assignment would be generated with more toll road itineraries selected as trip length increases. Concrete examples of this effect have been provided elsewhere (*13*).

Also discussed briefly is how these extensions can be incorporated into the standard mode choice and traffic assignment algorithms.

ACKNOWLEDGMENTS

This research was supported by a grant from l'Agence Française de l'Environment et de la Maîtrise de l'Energie.

REFERENCES

1. Gronau, R. Economic Approach to Value of Time and Transportation Choice. In *Transportation Research Record* 587, TRB, National Research Council, Washington, D.C., 1976, pp. 1–5.
2. Becker, G. S. A Theory of Allocation of Time. *Economic Journal*, Sept. 1965.
3. Abraham, C., and J. D. Blanchet. Le Modèle Prix-Temps, *Revue de l'Aviation Civile*, June 1973.
4. Marche, R. Valuation of Travel Time—French Intercity Values. Presented at Symposium on Transportation and Environment, Southampton, April 1973.
5. Pearman, D., J. Swanson, E. Kroes, and M. Bradley. *Stated Preference Techniques: A Guide to Practice*, 2nd ed. Hague Consulting Group, Den Haag, 1991.
6. Kanaan, A. *An Economic/Financial Analysis of Implementing Electronic Toll Collection and Traffic Management Systems*. M.S. thesis. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1992.
7. OEST. L'utilisation des véhicules de transport routier de marchandises en 1991, Ministère Français de l'Equipment, de Logement, et des Transports, Paris, 1992.
8. Ben-Akiva, M., D. Bolduc, and M. Bradley. Estimation of Travel Choice Models with Randomly Distributed Values of Time. In *Transportation Research Record 1413*. TRB, National Research Council, Washington, D.C., 1993, pp. 88–97.
9. Leurent, F. Equilibre Offre-Demande et Répartition entre Itinéraires dans les Modèles Statiques d'Affectation du Trafic. NT 92.2, INRETS working paper, INRETS, Arcueil, France, 1992.
10. Morellet, O. MATISSE: Un Modéle Intégrant Etroitement Induction et Partage Modal Fin du Trafic, Rapport INRETS #129. INRETS, Arcueil, France, 1990.
11. Sheffi, Y. *Urban Transportation Planning*. Prentice-Hall, Englewood Cliffs, N.J., 1985.
12. Wynter, L. M. The Value of Time of Freight Transport in France: Estimation of Continuously Distributed Values from a Stated Preference Survey, *International Journal of Transportation Economics*. In press.
13. Wynter, L.M. MIRAGE: A Traffic Assignment Model for Interurban Road Freight Transport in France. *Recherche, Transports, Sécurité*, Vol. 45, Dec. 1994, pp. 29–38.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Survey Methodology for Collecting Freight Truck and Destination Data

KENNETH L. CASAVANT, WILLIAM R. GILLIS, DOLLY BLANKENSHIP, AND CHARLES HOWARD, JR.

State departments of transportation and metropolitan planning organizations require a specific focus on freight and goods movements as one element of their planning process. A particular challenge is obtaining comprehensive information on freight truck movements. The Washington State Department of Transportation initiated a statewide freight truck origin and destination (O-D) study in April 1993 to meet this challenge. The Washington study is the first in the United States to collect statewide freight truck O-D data through direct personal interviews of truck drivers. Over 300 community service club members were hired and trained to conduct personal interviews at 28 separate locations throughout the state of Washington. A total of 30,000 truck drivers were interviewed, providing Washington with an extensive data base on statewide freight and goods movements. The methodology and procedures utilized to collect statewide freight truck data in Washington are described. Specific issues include research design, interview team recruitment and training, field data collection procedures, as well as ongoing project management requirements. Lessons learned from the Washington study provide insights for other states or regional planning organizations contemplating a freight truck study.

Providing for the efficient intermodal movement of freight and goods is a primary responsibility of state departments of transportation, metropolitan planning organizations (MPOs), and many local governments. This responsibility has received increased emphasis as a result of the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA), which requires states and MPOs to include a specific focus on freight and goods mobility as one element of their updated transportation plan.

Planning for the efficient movement of freight and goods is hindered by a lack of information on the source and characteristics of freight truck movements on state and regional highways. Freight movement by rail and water can be tracked adequately through Interstate Commerce Commission (ICC) waybill samples, the Corps of Engineers Waterborne Commerce data, and other published sources. However, obtaining comprehensive information on truck freight movements is much more difficult because of the large number of carriers and the numerous potential origins and destinations.

To address this information gap, the Washington State Department of Transportation (WSDOT) initiated a statewide freight truck origin and destination (O-D) study in April 1993. A regionwide freight truck O-D study was first proposed in Washington as an element of the Eastern Washington Intermodal Transportation Study (EWITS). EWITS is a 6-year ISTEA planning study to define the multimodal network necessary for the efficient movement of freight and people throughout the region of Washington on the east side of the Cascade Mountains. Supplemental funding provided by WSDOT enabled the EWITS freight truck O-D study to be expanded to include the entire state. Washington State University and The Gillis Group, a private consulting firm, were asked to conduct the study.

The Washington study is the first in the United States to collect statewide freight truck O-D data through direct personal interviews of truck drivers. The statewide study involved over 300 persons conducting personal interviews at 28 separate locations. A total of 30,000 truck drivers were interviewed to provide Washington with an extensive data base on statewide freight and goods movements.

The methodology and procedures utilized to collect statewide freight truck data in Washington are described. Specific issues include research design, interview team recruitment and training, field data collection procedures, as well as ongoing project management requirements. In the final section of this paper lessons learned from the Washington study are highlighted as insights for other states or regional planning organizations contemplating a freight truck study.

RESEARCH DESIGN ISSUES

Data Source Alternatives

Aggregate information on U.S. freight truck movements can be obtained from a variety of government and private-sector sources. For example, the Transearch freight flow data base compiled by Reebie and Associates provides aggregate information on commodity movements by mode between major cities. The Census Bureau and the Federal Highway Administration produce the Truck Inventory and Use Survey, the Nationwide Truck Activity and Commodity Survey, and the Commodity Transportation Survey. These government sources also provide a broad picture of major truck flows between regions. However, none of these sources is designed to provide information on freight truck movements in substate regions outside major cities or local transportation corridors.

The development of a methodology that would provide statistically reliable and comprehensive information on truck movements throughout the entire state was needed to fulfill research goals outlined for this Washington State freight truck O-D study. In particular, information on a wide array of freight truck characteristics is needed to plan effectively for the statewide freight and goods system. Examples include information on time-of-day movements, truck/trailer configuration, cargo type, payload weight, use of inter-

K. L. Casavant, Washington State University, Department of Agricultural Economics, 101 Hulbert Hall, Pullman, Wash. 99164-6210. W. R. Gillis and D. Blankenship, The Gillis Group, 108 North Adams, Ritzville, Wash. 99169. C. Howard, Jr., WSDOT Planning Department Transportation Building, P.O. Box 47370, Olympia, Wash. 98504-7370.

modal facilities, and routes utilized between major origins and destinations. Because this information is not available from published secondary sources, it is necessary to collect data directly from trucking firms.

Several specific criteria were developed as guidelines for the design of the statewide truck survey project:

- Data collected should provide statistically reliable information on truck characteristics and commodity flows for all major Washington highways.
- The sample size should be large enough to provide useful freight and goods movement information for major transportation planning regions as well as the state as a whole.
- Information should be developed over a continuous 24-hr period during each of the four seasons.

Among the alternatives, including mail or telephone surveys, roadside interviews of truck drivers are the most effective means of generating truck freight information addressing these three criteria. Several previous studies utilizing roadside interviews provided insights into the development of the Washington freight truck O-D study. Among the most comprehensive of the previous studies is the Ontario Commercial Vehicle Survey of 1988. The Ontario Ministry of Transportation conducted roadside surveys of 19,225 commercial vehicles between March and November 1988. The interviews were conducted by Ministry of Transportation staff at 41 weigh stations and 16 additional roadside locations. The truck driver interviews collected a wide array of information ranging from axle counts and cargo type to vehicle weight. A similar study conducted by Washington State University in 1992 focused on the northwestern Washington border crossing between Canada and the United States. Roadside interviews were conducted by student interviewers as trucks passed through the U.S. port of entry. In addition to collecting information on truck characteristics and commodity type, the Washington State University study documented specific highway routes utilized by the trucks.

Site Selection

Following the lead of the previous studies, permanent weigh stations and ports of entry were utilized as the primary data collection sites for the Washington truck survey. The specific weigh stations utilized as data collection sites were identified through analysis of WSDOT's traffic count and vehicle classification data. Data collection sites were established on all state highways with a significant volume of daily truck traffic. On Washington's major Interstate corridors, multiple data collection sites were identified. In recognition of the importance of expanding international trade, plans were also established for roadside interviews at major U.S.-Canadian border crossings.

To obtain a complete profile of truck movements it is necessary to interview truck drivers traveling in both directions on a given highway segment. On divided highways, this requires identifying two separate interview sites on alternate sides of the highway. In total, interview sites were established at 21 Washington State Patrol weigh stations, three Canadian border locations, and the Oregon port of entry in Umatilla (Figure 1).

Questionnaire Design

Questionnaire design is an important element of a successful methodology for roadside truck driver interviews. The Washington

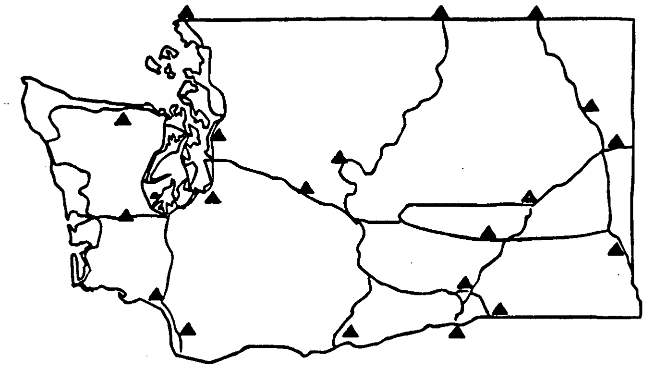


FIGURE 1 Truck interview locations.

statewide truck driver survey collected information on time-of-day movements, vehicle configuration, trucking company location, origin and intended destination, cargo type, vehicle and cargo weight, use of intermodal facilities, and the specific route traveled (Figure 2).

To encourage truck driver participation, the questionnaire was designed to be completed within 3 min. Approximately one-half of the questions (for example, number of axles, trailer style, time of day, hazardous material placard) could be filled out through direct observation by trained interview personnel. Questions to be asked directly to the truck driver focused on cargo, weight, use of intermodal facilities, and route of travel. To the extent possible, check boxes were utilized to minimize necessary writing and enable rapid completion of each interview. A map of major Washington State highways was attached to each questionnaire. Utilizing this map, the interviewers were able to quickly highlight Washington highways utilized by drivers traveling between their stated origin and destination.

Appropriate phrasing of interview questions is also essential. For example, the preliminary questionnaire developed for the Washington truck driver survey requested information on the "payload weight." During the pretest phase of the project, many drivers confused "payload weight" with "gross weight." Consequently, the questionnaire was revised to request, "What is the weight of the cargo that you are carrying today?" Establishing a process for ongoing evaluation and modification to the survey questionnaire is essential.

Interview Dates and Duration

The Washington freight truck O-D study developed data for each of the four seasons—summer, fall, winter, and spring. Particularly in agricultural regions, seasonal differences in truck movements can be substantial. For example, fresh potatoes and grains are harvested in the summer. Apples are harvested in the fall. Consequently, highways important to the transportation of these commodities may have significantly different truck volumes in different seasons.

The Washington truck survey was designed to provide a profile of statewide truck movements during each season. Consequently, it was important to conduct interviews at all 25 sites within as short a time frame as possible during each season. Data collection sites were systematically scheduled to avoid the practical problem of requesting multiple interviews from the same truck driver on a given day. With this constraint, approximately 5 weeks was needed to complete each interview cycle at the 25 sites.

Truck driver interviews at most locations were scheduled for a continuous 24-hr period to provide a comprehensive picture of statewide truck movements. Interviews were also consistently scheduled for Wednesdays to obtain median traffic patterns rather than exceptionally heavy Monday or Friday flows. Maintaining consistency in day-of-week data collection at each site helps to avoid potential statistical bias when the data are aggregated to profile statewide movements.

Sampling Issues

Approximately 4,500 commercial vehicles travel Washington's I-5 corridor each day. Other major Washington freight truck corridors typically support 1,500 to 2,500 trucks per day. Interviewing every truck driver traveling these busy corridors is neither feasible nor necessary. A systematic sampling strategy was developed for the Washington truck survey.

An overall goal of obtaining at least 300 surveys over a 24-hr period at each site was established. One out of every 10 trucks on I-5 and 1 out of every 5 trucks on most other major Washington freight truck corridors were targeted for an interview. On several of the lower-volume routes, 1 out of every 2 commercial vehicles was targeted for an interview. A total of approximately 7,500 truck drivers were interviewed during each of the four survey rounds, providing a data base of 30,000 interviews for the year-long study.

INTERVIEW TEAM RECRUITMENT TRAINING

Recruitment

To obtain an accurate seasonal profile of truck movements throughout the state of Washington, it was necessary to conduct interviews simultaneously at more than six sites across the state. The more typical approach of hiring a team of interviewers to travel from site to site over a period of weeks did not meet the research design goals established for the Washington study. Typically, 15 people are required to cover a 24-hr interview session at each of the sites. On a given day, up to 90 interview personnel are required. Consequently, it was a particular challenge to obtain a very large short-term labor force to successfully complete the study.

To meet this challenge, members of community service clubs with statewide membership were recruited and trained to conduct truck driver interviews. The opportunity was first introduced at a statewide conference of Washington Lions Clubs, which were offered the opportunity to conduct interviews as a fund-raising activity.

As a result of initial recruitment efforts, 15 Lions Clubs and 1 Kiwanis Club agreed to provide at least 15 members who would serve as a local interview team. A total of over 300 service club members participated in the Washington study.

All the clubs were located close to the selected interview sites, which minimized travel costs for the interview team. The club members' personal knowledge of local roads and industries also proved to be a major advantage in communicating and understanding responses provided by truck drivers. Most of the same club members participated in each of the four interview rounds, creating a highly experienced local interview team for future projects.

Training

Interview team training is always an important component of any study involving personal interviews, and a strong training program is essential when less-experienced personnel are used. A detailed training program for interview team members, including both classroom and on-site instruction, was conducted by The Gillis Group:

Topic	*Time Allotted (min)*
Project goals and objectives	10
Overview of interview questionnaire	20
Identifying truck and trailer configurations	15
Personal interview techniques	15
Safety requirements	10
Things to bring	5
Questions and answers	30

An individual and customized classroom training session was conducted for each of the 16 service clubs. Each training session began with an overview of the key project goals and objectives so that interviewers would be prepared to answer basic questions from truck drivers concerning the purpose of the study. The interview questionnaire was reviewed in detail. Particular focus was given to ensuring that the interview team members were able to accurately identify the different truck and trailer configurations. Personal interview techniques were also covered. In particular, advice was offered on how to introduce the purpose of the study and request an interview from the truck driver.

Conducting personal interviews of truck drivers at busy weigh stations is a strenuous and potentially dangerous activity. Every effort was made to design a site setup and traffic control plan to prevent the possibility that an unwary interviewer might step into the path of an oncoming truck. In addition, the personal responsibility of each interviewer to be alert and promote on-site safety was stressed in the training. Examples of safety requirements emphasized to the interview team members were to (a) always wear safety vests and hats while on site, (b) never approach a truck when it is moving, (c) not allow traffic congestion to occur in the interview area, and (d) take regular breaks. A manual outlining safety requirements, truck configurations, and other interview guidelines was provided to each team member at the conclusion of the classroom training.

The classroom training session is only the beginning of what should be a continuous process to ensure quality interviews and personnel safety. Ongoing training and instruction were provided by a supervisor assigned from the project management team to each site. Over time most teams became highly adept at conducting the personal interviews, and constant supervision was no longer necessary. However, periodic monitoring of interview activities continued throughout the project.

FIELD DATA COLLECTION PROCEDURES

Equipment Needs

Proper equipment is necessary to ensure both interview personnel safety and accurate results. Equipment utilized at each interview site is as follows: reflective safety vests; clipboards; survey crew signs; headlamps; interview team hats; pens, pencils, and highlighters; weatherproof boxes; and traffic cones. An adequate supply of basic equipment such as clipboards, pens, pencils, highlighters, and staplers is necessary. A large Survey Crew sign was utilized at each site both to inform truck drivers approaching the station that an official survey was taking place and to caution that interview personnel were in the vicinity. All interview personnel were issued safety vests. It is important that high-quality reflective vests be utilized when interview crews are operating in the dark. Headlamps were

For Office Use Only
Survey________
QCA________
Input ☐

WASHINGTON STATE DEPARTMENT OF TRANSPORTATION AND
WASHINGTON STATE UNIVERSITY
TRUCK TRAFFIC SURVEY, SPRING 1994

Please Remember - your Club is depending on YOU for the Quality Control Award!

✓ *Write neatly!* ✓ *Do not abbreviate!* ✓ *Complete all required questions!*

Thank You!

CONFIDENTIAL

1) Station location: Tokio Westbound

2) Initials of interviewer: ____________________

3) Interview shift:

1. Day Shift	2. Evening Shift	3. Night Shift
7:00 a.m. - 3:00 p.m.	3:00 p.m. - 11:00 p.m.	11:00 p.m. - 7:00 a.m.

4) Time of interview: ______ a.m. ______ p.m.

5) Is this truck a part of the "official sample"? a) ☐ Yes b) ☐ No

6) Truck Configuration	7) Trailer Style
[Check only one truck configuration]	*[If appropriate, check more than one trailer style]*
[See Quality Control Notes for definitions]	***[See Quality Control Notes for definitions]***
1. ☐ Straight truck	1. ☐ Van (without temperature control)
2. ☐ Truck and trailer	2. ☐ Van with temperature control
3. ☐ Tractor only	3. ☐ Flatbed
4. ☐ Tractor and trailer	4. ☐ Car carrier
5. ☐ Tractor with two trailers	5. ☐ Hopper or belly dump
6. ☐ Other (specify) __________	6. ☐ Stake and rack
	7. ☐ Concrete mixer
	8. ☐ Tanker
	9. ☐ Float or low boy
	10. ☐ Dump
	11. ☐ Container
	12. ☐ Wood chip
	13. ☐ Animal carrier
	14. ☐ Logging
	15. ☐ Belt
	16. ☐ Other (specify) ________

8) Total number of axles on the ground: __________

9) Is a hazardous material placard displayed? 1) ☐ Yes ID# __________ 2) ☐ No

FIGURE 2 Survey questionnaire.

[Please ask the following questions] **CONFIDENTIAL**

10) Trucking company name: ______________________________

Very Important! Do Not Abbreviate! Be Exact!

11) Trucking company home base: City: ______________ State/Province: ______________

12) What is the unloaded weight of this vehicle? ____________lbs.

13) Is this vehicle carrying cargo or is it empty? □ carrying cargo *[Ask Q14-21]* □ empty *[Ask Q22-27]*

14) What is the major commodity on board: ______________________________

Do Not Abbreviate! Be Specific!

15) How much does the cargo you are carrying today weigh? ____________ lbs.

Complete only the one column that applies to this trip. No round-trip information, please!

Trucks CARRYING Cargo:	Trucks WITHOUT Cargo:
Where did you pick up this cargo?	Where did you pick up this cargo?
16) City: ______________	22) City: ______________
17) State/Province: ______________	23) State/Province: ______________
18) Facility: *[see Quality Control Notes]*	24) Facility: *[see Quality Control Notes]*
a) □ trucking yard	a) □ trucking yard
b) □ railroad yard	b) □ railroad yard
c) □ river or ocean port	c) □ river or ocean port
d) □ airport	d) □ airport
e) □ factory, processing plant, or sawmill	e) □ factory, processing plant, or sawmill
f) □ warehouse/distribution center or post office	f) □ warehouse/distribution center or post office
g) □ farm or forest	g) □ farm or forest
h) □ retail store or gas station	h) □ retail store or gas station
i) □ job or construction site	i) □ job or construction site
j) □ other: ______________	j) □ other: ______________
What is the destination of your cargo?	Where will your trip without cargo end?
19) City: ______________	25) City: ______________
20) State/Province: ______________	26) State/Province: ______________
21) Facility: *[see Quality Control Notes]*	27) Facility: *[see Quality Control Notes]*
a) □ trucking yard	a) □ trucking yard
b) □ railroad yard	b) □ railroad yard
c) □ river or ocean port	c) □ river or ocean port
d) □ airport	d) □ airport
e) □ factory, processing plant, or sawmill	e) □ factory, processing plant, or sawmill
f) □ warehouse/distribution center or post office	f) □ warehouse/distribution center or post office
g) □ farm or forest	g) □ farm or forest
h) □ retail store or gas station	h) □ retail store or gas station
i) □ job or construction site	i) □ job or construction site
k) □ other: ______________	k) □ other: ______________

28) What Washington highways were used to travel between the two locations identified above?

______________________________ *(Remember, accurately highlight attached map!)*

Write out the highways used to get between the two locations identified above.

29) *Including this trip,* how many times has this truck traveled the above route in the past 7 days?

__________ Times □ Don't know

FIGURE 2 *(continued)*

also provided to the night shift to help interview team members see and be seen. Weatherproof boxes were provided at each site as storage for completed surveys.

Site Setup and Traffic Control

Typical site setup and traffic control plans utilized at interview sites are shown in Figures 3 and 4. Procedures at both large and small sites were similar. Site setup and traffic control plans at U.S./Canadian border locations were similar to those utilized at weigh stations.

All interview scheduling was coordinated and approved by the Washington State Commercial Vehicle Enforcement Division or U.S. and Canadian Customs officials in the case of data collection sites at border crossings. Cooperation from the Oregon Department of Transportation was received in conducting interviews at the Umatilla port of entry.

Uniformed officers at the weigh stations or ports of entry conducted enforcement activities as normal. Commercial vehicles entered the weigh station or port of entry checkpoints following the usual procedures. After enforcement activities were complete, selected trucks would be directed into a designated interview area by the officer in charge or a member of the interview team. Truck drivers were selected for an interview according to predetermined interview procedures. For example, along I-5 every tenth truck was directed to the interview area. Along I-90, every fifth truck was directed to the interview area. At some low-volume locations, every other truck driver was selected for an interview.

While the truck was parking, a member of the interview team completed visual information such as the time of day, number of axles, truck configuration, and the presence of a hazardous materials placard. After the truck came to a complete stop, a member of the team approached the driver requesting an interview. Truck driver participation in the survey was voluntary. Approximately 95 percent of the truck drivers requested to complete an interview agreed to participate. The truck drivers were provided with a coupon for a free cup of coffee as a token of thanks for their participation.

At the smaller weigh station sites, only two and at most three trucks can be safely parked to the side. At stations with a large park-

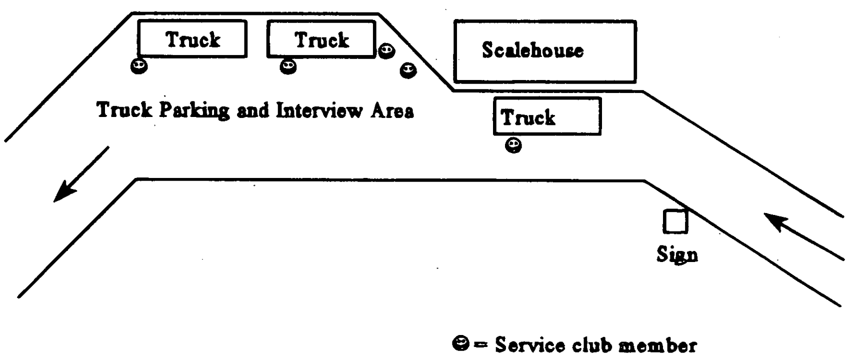


FIGURE 3 Site setup and traffic control plan at smaller weigh stations typical on lower-traffic-volume corridors.

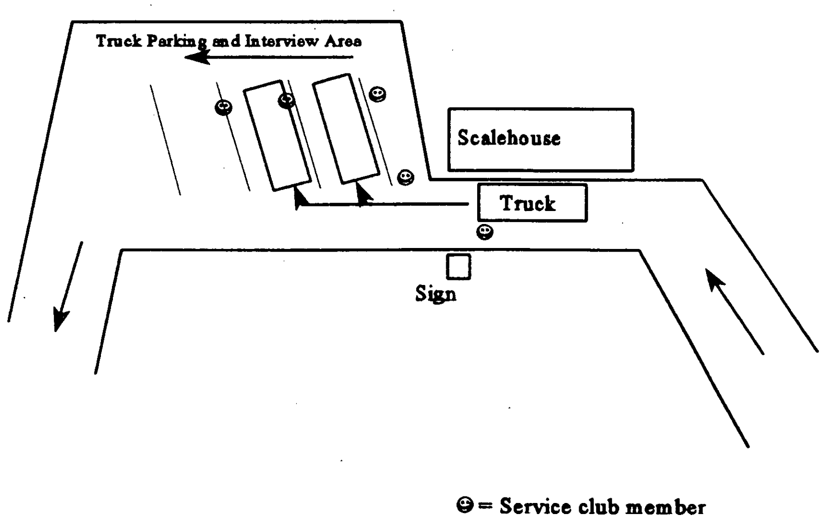


FIGURE 4 Site setup and traffic control plan at weigh stations with large parking area typical of higher-traffic-volume locations on I-5 and at state ports of entry.

ing area, traffic cones were utilized to block off up to four lanes as a designated truck interview area. In all cases a lane was provided to allow trucks not selected for an interview to pass safely back onto the highway.

On average, an experienced interviewer can complete the questionnaire within 3 min. A crew of up to five interview personnel was maintained at each site throughout the 24-hr survey period. This enabled the crew to quickly begin a new interview as soon as the previous one was completed. In most cases space and personnel at the interview sites were adequate to ensure a free flow of traffic. However, during certain busy periods trucks whose drivers would have otherwise been interviewed were allowed to pass by because space to park them safely was not available.

Importance of Uniformed Officers

Cooperation and support from Commercial Vehicle Enforcement Division officers and Customs officials were essential to the success of the field data collection project. Uniformed officers provided two critical services. First, their presence helped to ensure the safety of the interview personnel. Second, the presence of a uniformed officer was likely a major contributing factor to the high level of participation received from truck drivers asked to complete an interview. While passing through weigh stations or ports of entry, truck drivers are prepared to present their records and respond to questions pertaining to enforcement. Many of the truck drivers indicated that they thought they were being asked to park for an enforcement violation. They expressed relief and willing cooperation when they learned that they were only being asked some questions about their destination and cargo.

Quality-Control Procedures

A program of on-site quality control is essential to ensuring accurate data results from the interviews. A strong training program for the interview teams was one tool utilized to support accurate data collection. In addition, a supervisor from the project management team was assigned to each site to check questionnaires for accuracy as they were completed. Problem areas were immediately addressed with interview personnel as necessary.

A quality-control award was established as an extra incentive for service clubs to perform quality work. Small bonus checks were awarded to clubs that provided data entry personnel with the most legible, accurate, and complete questionnaires. As a result many of the clubs instituted their own quality-control measures. Many of the clubs assigned one individual to check each completed questionnaire and make changes as necessary. Several of the clubs rewrote questionnaires that were less legible.

Data quality was also affected by the weather and other events beyond the control of the team. For example, some paper questionnaires got wet during rain showers. Several of the interview sessions took place during snowstorms. To protect the safety of the crew and ensure high-quality data, interviews were often conducted inside the scale house during periods of bad weather.

Particularly along high-traffic-volume corridors, brief interruptions to the data collection occurred. For example, during high-volume periods, the available capacity on the weigh station entry ramp may be inadequate. In these cases, to prevent congestion on the highway it was necessary to shut down enforcement and interview activities until the traffic could clear. In several cases road construction in the vicinity of the weigh station made it necessary to cease interview operations for a period of several hours. At one station, a hazardous material spill resulted in the closing of an interview site for an entire afternoon. Interview teams recorded any breaks in activity or other problems affecting data quality on a site summary sheet provided to the project managers.

LESSONS FROM WASHINGTON FREIGHT TRUCK STUDY

The Washington study is the first in the United States to collect statewide freight truck O-D data through direct personal interviews of truck drivers. This final section highlights several lessons learned from the Washington study as insights for other states or regional planning organizations contemplating a freight truck study.

Lesson 1: Community service clubs can be a viable labor force for conducting personal interviews of truck drivers. However, strong management systems are required. Over 300 community service club members participated as interviewers for the study. In general, service club members proved to be an able and effective labor force for conducting truck interviews. By utilizing service clubs it was possible to assemble a very large labor force willing to work for a 1-day period on a scheduled date. The ability to assemble workers living close to selected interview sites was also an advantage. Individuals living in the area often had personal knowledge of roads and local industries that proved helpful in communicating and accurately recording responses offered by truck drivers.

The use of service clubs as a labor force provided additional benefits to the state and regions beyond data collected for the study. Funds raised from this project were used by the clubs to provide services in their communities, for example, Little League baseball teams, construction of an outdoor community amphitheater, upgrading of bleachers at the baseball field, support of local hospices, provision of eyeglasses and hearing aids to the elderly and disadvantaged, corneal transplants, and ongoing medical research. In addition, hundreds of Washington residents gained a new appreciation of the importance of the highway system in transporting products that they utilize in their daily lives.

Service clubs offer many advantages as an interview labor force. However, they also present several critical challenges that must be addressed through strong management systems. The most significant of these challenges is a lack of uniform skills and physical abilities among service club members. Implementing a strong and consistent training program is essential when utilizing service club members, who may have no previous experience in conducting personal interviews. Training should include both classroom instruction as well as on-site supervision.

The majority of service club members are highly capable individuals with excellent skills required to accurately record information provided by truck drivers. However, the personal skills or physical condition of some members can be a significant hindrance to quality data collection. Consequently, it is recommended that written agreements with the clubs be developed to ensure that persons unable to perform quality work are not assigned to interview tasks. Frequent project management evaluations are recommended to help identify individuals with particular problems. When possible, additional assistance may be offered to help correct data collection problems associated with a particular individual. However, the project management team should retain the authority to dismiss an individual from the interview team if necessary.

Lesson 2: Involvement of uniformed enforcement officers is a critical factor in obtaining cooperation and participation from truck drivers requested to complete interviews. A remarkably high level of cooperation and participation from drivers was obtained in the study. Even though participation was voluntary, approximately 95 percent of the truck drivers asked to participate agreed to an interview. The constant presence of a uniformed officer at all sites was a major factor contributing to participation. When a truck driver interview study is planned, it is essential to solicit cooperation from Commercial Vehicle Enforcement Division officers and Customs officials (for border crossings). In the state of Washington, these organizations were enthusiastic partners. Results from the statewide truck study will also be helpful in planning for more effective commercial vehicle enforcement activities.

Lesson 3: Site setup and the use of systematic sampling techniques are important factors to maintain traffic flows and promote cooperation at the interview sites. The approach of systematically selecting trucks for an interview according to their position in line as they entered the weigh station proved to be an excellent technique to promote smooth traffic flows at the data collection sites. A rhythm was established in which one truck was able to pull out of the designated interview area as another truck pulled in. A commitment to keep personal interviews under 3 min encouraged cooperation among the truck drivers, many of whom were on a tight schedule but willing to spend 3 min. Every station is unique. Enough interviewers were present on site to ensure that drivers did not have to wait for an interview. To maintain interview team safety and cooperation from the drivers, careful attention to site setup and sampling procedures is essential.

Lesson 4: Establishing ongoing procedures for evaluation and modification of procedures is important to quality data collection. Establishing formal systems for evaluation and improvement of data collection procedures is a key element of a successful survey methodology. Members of the project management team maintained on-call accessibility to both the Washington State Patrol and the clubs during the interviews. An individual from each club was identified as survey coordinator, whose primary responsibility was to coordinate and communicate all survey-related activities among club members. Weekly phone calls were made to both the Washington State Patrol and the service club coordinators to remind them of upcoming survey dates. This consistent communication worked well to build rapport and credibility. It also provided everyone with an opportunity to discuss any issues as they arose and to stay current with the project.

After the second round of data collection, a more formal evaluation took place. Telephone interviews were conducted with each Commercial Vehicle Enforcement officer and with each club coordinator. The interviews provided the opportunity for feedback on the interview process from both a safety and a field perspective. This evaluation process resulted in several changes to the survey instrument to make questions clearer. It also provided a structured review of the data collection process and suggestions for improvements. Most important, the ongoing evaluation process helped to maximize the quality of data collected.

ACKNOWLEDGMENTS

This work was funded by the Washington State Department of Transportation, Washington State University Eastern Washington Intermodal Transportation Project, and the Intermodal Surface Transportation Efficiency Act of 1991.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Design and Implementation of a Statewide Roadside Origin-Destination Survey in Vermont

UDAY VIRKUD AND CORINE S. KEYES

The results are summarized of an origin-destination survey conducted to estimate external-to-external and external-to-internal trip tables of a statewide travel demand forecasting model. Costs and effectiveness of the interview technique and the mail-back postcard technique for collecting these data are compared, and the following topics are discussed: roadway selection, survey design, sample size estimation, field data collection, survey results, and survey costs. The results of the study indicate that the interview technique is appropriate for low-volume roadways, whereas the mail-back postcard technique is not cost-effective in these low-volume situations. However, in high-traffic-volume situations, with a good traffic control plan and a well-trained crew, the mail-back postcard technique can be implemented safely with no substantial traffic delays.

The Vermont Agency of Transportation (VAOT) and its consultant, Vanasse Hangen Brustlin, Inc., (VHB), in their efforts to develop a statewide travel demand forecasting model, decided to collect statewide origin-destination (O-D) data for the first time in Vermont. The process of planning, designing, and implementing this unique large-scale data collection effort is described.

Vermont is one of the smallest states in the country in terms of size. It borders New York, New Hampshire, Massachusetts, and Canada. There are approximately 70 roads that can be used to enter Vermont; 25 were selected for roadside O-D surveys. Traffic volumes on these 25 roads represented approximately 90 percent of all traffic entering Vermont.

The most important cordon line information for the study area necessary for model development is the vehicle O-D data. It is also important to gather information regarding the time of day during which the trip was made and the purpose of the trip. This information provides data necessary to synthesize external-to-external trip tables and trip-length frequencies for calibrating external-to-internal travel patterns. The time of travel can also be used to estimate peak period travel from the daily traffic forecasts, and the trip purpose provides information necessary for developing travel demand management strategies. Vehicle type and vehicle occupancy data are also collected.

Two survey techniques, a mail-back survey and an interview survey, were selected for this study. The performance and costs of the two techniques are compared and the steps involved in site selection, selection of survey techniques, sample size estimation, questionnaire design, traffic control, and safety measures are described. Transportation planners should find the description of the data collection effort and the costs associated with it to be helpful because there has been a renewed interest in such efforts created by recent legislative changes such as the Intermodal Surface Transportation Efficiency Act of 1991 and the Clean Air Act Amendments of 1990.

Vanasse Hangen Brustlin, Inc., 101 Walnut Street, P.O. Box 9151, Watertown, Mass. 02272.

ROADWAY SELECTION

To obtain a representative sample of one-way, inbound traffic entering Vermont, 25 roadways crossing the state line were selected for collecting O-D data. The selection of roadways for the survey was based on three primary factors: geographic location, functional classification, and average annual traffic volume. A good sample of statewide O-D data should include different types of roads that are located throughout the state and that represent a wide range of traffic volumes. All Interstates and principal arterials were selected, in addition to half of the minor arterials and one-third of the major collectors. The average annual daily traffic (AADT) along the selected roadways equal approximately 155,000 vehicles, representing 90 percent of total cordon line traffic.

At 20 of the 25 selected locations, surveys were conducted by VHB staff. The other 5 locations were surveyed by the New Hampshire Department of Transportation (NHDOT), which is in the process of collecting similar data for its planning studies. The two states cooperated on this effort to save funds and improve the compatibility of the two states' modeling efforts. The data for the five locations surveyed by NHDOT are not included in this paper.

The data collected at the selected roadways are being analyzed to develop trip tables. During the analysis, the sample data will be weighted to represent daily traffic volumes at each station, since only the busiest 12-hr period was sampled. The travel patterns along the roadways where O-D data were not collected will be estimated on the basis of O-D patterns for similar roadways in the vicinity.

SURVEY DESIGN

Survey design includes selection of appropriate survey techniques and adequate sample sizes for surveying, determination of the information to be collected, design of the questionnaires, selection of survey stations, and scheduling of the survey.

Survey Technique

Various O-D survey techniques were evaluated for possible use, including roadside interviews, mail-back postcard surveys, license plate trace surveys, license plate mail-out surveys, and tag-on-vehicle/lights-on surveys (*1*).

The license plate trace method and the tag-on-vehicle/lights-on surveys are suitable for small study areas because vehicles can be tracked over short distances. However, vehicles cannot be tracked over an entire state. In addition, these types of surveys do not produce information about the trip purpose, which is important for model development.

The license plate mail-out survey involves recording license plate numbers of vehicles on a selected roadway, tracing vehicle ownership, and mailing a survey to owners. Because this method does not require vehicles to be stopped to receive the survey, it is less disruptive to traffic flow. However, because of the large number of out-of-state vehicles expected to be traveling into Vermont and the amount of work involved in tracing their ownership, this method was not considered appropriate. Furthermore, since the surveys are mailed at a later date to the owner of the vehicle (and not necessarily to the motorist who made the trip on the survey day), the accuracy of the data from this method is expected to be lower than that of the roadside postcard survey.

The roadside interview involves directing vehicles into a designated interview area and asking a series of short questions. The advantages of this technique are that it provides more complete information than other techniques and has a higher response rate. The disadvantage is that it generally requires more personnel and traffic control measures than other techniques.

Postcard surveys are distributed to motorists either at a location where they normally stop or after they are brought to a stop on a roadway. The advantages of this technique are that postcards can be distributed quickly and with fewer personnel than are required for interviews. The disadvantage is that a higher number of vehicles must be sampled to obtain an adequate number of completed surveys because the typical response rate for mail-back surveys is generally between 15 and 30 percent. The postcard technique is generally suitable for higher-volume roads where conducting interviews could cause longer traffic delays or at locations where it is not feasible to conduct interviews.

The roadside interview and the postcard survey were selected for the Vermont study because both methods are appropriate for large study areas and both can be adapted to collect information needed for the traffic model, including trip purpose. The survey method chosen for a particular site was based on several factors, including traffic volume, physical constraints, language needs along the Canadian border, and restrictions set by the U.S. Customs and Immigration offices at the Canadian border sites. In addition, the sample size requirements described in the following section were an important criterion for determining the survey method at each site. Because of lower response rates, the mail-back postcard surveys generally are not able to obtain adequate sample sizes on low-volume roads. Interviews can be conducted along low-volume roadways to obtain the necessary sample size and at other sites where the surveys are not expected to have an adverse impact on traffic operations.

Sample Size

Generally, O-D surveys rely on one of two sampling approaches (2). The first approach focuses on providing a sufficient sample size for direct estimation of traffic flow at an aggregate (district-to-district) level. The second approach focuses on the statistical requirements of estimating a proportion as it relates to a variable of interest, such as average trip length or trip purpose. Neither approach correlates explicitly with the primary purpose of the O-D survey—estimation of external trip tables.

For the purposes of this study, two requirements were established. The first was that the survey should result in a reasonably accurate external-to-external trip table; the second was that it should provide average trip length data for calibration of the external-to-internal trips. The sampling rates for the first approach mentioned above are generally higher, and therefore critical. As a result, the first approach was selected to meet the sample size requirements for the purposes of this study.

The required sample size for each survey station is defined as the minimum number of usable postcards or interviews completed. The sample size is usually represented as a rate that is the ratio of the total number of vehicles sampled to the total number of vehicles passing through. The sample size formula is given by (*3,4*)

$$r = (Z^2pq)/[(N - 1)\,W^2 + (Z^2pq)]$$

where

r = sampling rate,
p = proportion of total traffic volume at the survey station that has a particular O-D pair,
$q = (1 - p)$,
W = desired accuracy: (percent error $\times\ p$),
N = traffic volume at the survey station, and
Z = normal variate that is associated with a specified level of confidence in estimating the O-D interchange volume.

For the purpose of this study the desirable sample sizes were estimated at 90 percent confidence, errors within ± 15 percent, and $p = 15$ percent or lower. A sample calculation for a roadway with a one-way, 12-hr volume equal to 1,300 vehicles follows:

$p = 0.15$
$q = 0.85$
W = percent error $\times\ p = 0.15 \times 0.15$
$N = 1{,}300$
$Z = 1.645$, the normal variate for 90 percent confidence
$r = (2.706 \times 0.15 \times 0.85)/[(1299 \times 0.0005) + (2.706 \times 0.15 \times 0.85)]$
$r = 0.34$

Because sample sizes were calculated on the basis of historical traffic volumes and because of the fluctuations in day-to-day traffic, sampling rates were estimated conservatively for a range of traffic volumes and not for individual roadways. The following sampling rates were determined to meet the sample size criteria presented above: a sampling rate of 0.34 for roadways with traffic volumes under 5,000 vehicles during the 12-hr survey period; a rate of 0.24 for roads with volumes between 5,000 and 10,000 vehicles during the 12-hr survey period; and a rate of 0.14 for roads with volumes between 10,000 and 20,000 vehicles during the 12-hr survey period. For roads with volumes over 20,000 vehicles, a lower rate could be used; however, there were no roads in Vermont in this category.

The estimated number of postcards to be distributed and interviews to be conducted at each site are presented in Table 1. The AADTs used to estimate the sample sizes were obtained from his-

TABLE 1 Estimated Required O-D Sample Sizes and Survey Types

Route Number	Functional Classification	Two-Way AADT*	Estimated One-Way 12 Hr. Volume**	Sample Rate	Sample Size	Required Surveys Mail-back x0.30	Required Surveys Interview x0.95	Survey Type
US 2	Principal Arterial	3,386	1,270	0.34	432	-	455	Interview
VT 314	Major Collector	1,727	648	0.34	220	-	232	Interview
VT 17	Minor Arterial	2,075	778	0.34	265	-	278	Interview
US 4	Principal Arterial	7,634	2,863	0.34	973	3,245	-	Mail-Back
VT 149	Major Collector	2,787	1,045	0.34	355	-	374	Interview
VT 9	Principal Arterial	9,130	3,424	0.34	1,164	3,881	-	Mail-Back
VT 114	Major Collector	3,065	1,149	0.34	391	-	411	Interview
US 2	Minor Arterial	4,138	1,552	0.34	528	-	555	Interview
VT 10A	Major Collector	14,414	5,405	0.24	1,297	4,324	-	Mail-Back
NH 119	Minor Arterial	6,700	2,513	0.34	854	2,848	-	Mail-Back
NH 25	Major Collector	3,000	1,125	0.34	383	1,275	-	Mail-Back
VT 123	Minor Arterial	4,500	1,688	0.34	574	-	604	Interview
US 7	Principal Arterial	4,956	1,859	0.34	632	2,107	-	Mail-Back
VT 100	Major Collector	2,878	1,079	0.34	367	-	386	Interview
I 91	Interstate	14,471	5,427	0.24	1,302	4,342	-	Mail-Back
VT 142	Major Collector	1,168	438	0.34	149	-	157	Interview
I 81/US 7	Interstate	2,929	1,098	0.34	373	1,244	-	Mail-Back
VT 139	Major Collector	1,215	456	0.34	155	517	-	Mail-Back
I 91/US 5	Interstate	5,069	1,901	0.34	646	2,154	-	Mail-Back
VT 147	Major Collector	1,117	419	0.34	142	475	-	Mail-Back
Total					11,203	26,411	3,453	

* AADT based on previous years' traffic counts provided by VAOT.
** Estimated 12-hour volumes are equal to two-way AADT divided by two, then multiplied by 0.70 to obtain a 12-hour estimate. This number was chosen based on an analysis that showed that approximately 70 percent of AADT volumes were present on the roadways between 7:00 AM and 7:00 PM.

toric traffic volume data provided by VAOT. The minimum number of usable surveys required for each location was estimated on the basis of the sampling rates described above. A return rate of 30 percent was assumed for postcard surveys and a response rate of 95 percent was assumed for interview surveys.

Survey Questionnaires

The mail-back postcards and the interview forms were designed to gather specific information about the one-way, inbound trips, including:

- Trip origin, including street address, city or town, and zip code;
- Trip destination, including street address, city or town, and zip code;
- Trip purpose;
- Vehicle occupancy; and
- Vehicle type.

This information will provide sufficient detail to identify O-D traffic analysis zones and develop the trip tables for the model. However, two formats are required because of the difference in the way the questions are asked of motorists and in the way responses are recorded.

Postcards

The mail-back postcards provided precise instructions to respondents in order to gather the required information about the one-way trip. Since respondents had no chance to ask questions about proper recording of answers on the postcards, the instructions had to be as clear and concise as possible to reduce error.

The postcards were divided into a question portion and an answer portion. The answer portion was detachable and had the address and prepaid postal imprint on the back. This allowed adequate space to provide instructions while still meeting U.S. Postal Service requirements to minimize postage costs. Postcards in two languages (English and French-Canadian) were distributed at sites along the Canadian border for those motorists unable to respond in English. Each postcard was individually numbered to determine the time and location of distribution. A sample postcard is shown in Figure 1.

Interview Forms

The interview forms consisted of a set of questions with instructions to the interviewers and interview response sheets. The instructions guided the interviewers through the questions and reminded them of the level of detail needed in the responses. In addition, a separate detailed project description sheet, intended to be handed to any

Travel Survey ▪ Vermont Agency of Transportation and New Hampshire Department of Transportation

Questions

Please take a moment to answer a few questions about your trip into or through Vermont (**excluding** any return trip). Your responses will be used to help determine the need for transportation improvements in this area. **Please record your responses on the attached card ▶**

1 Where did your trip begin? (the **last** place you entered your vehicle prior to receiving this card, excluding short stops for gas or food) 1

2 What type of place is your trip start point? 2

3 Where did your trip end? (the **first** place you exited your vehicle after receiving this card, excluding short stops for gas or food) 3

4 What type of place is your trip end point? 4

5 If your trip ended outside of Vermont, please specify which route you used to leave the state (check one) 5

6 What was the purpose of your trip? 6

7 How many people were in the vehicle, **including** the driver? 7

8 What type of vehicle were you in? 8

9 Please add any comments on transportation you may have. 9

Please complete, detach, and return the answer portion of the postcard as soon as possible. No postage is necessary. **Thank you very much for your cooperation!**

Travel Survey ▪ Vermont Agency of Transportation and New Hampshire Department of Transportation

Answers

Information should be provided only for the trip you were making when you received this card.

1 Street Address ______
Nearest Intersection/Landmark ______
Town ______ State/Province ______ Zip ______

2 ❑ Your Primary Residence ❑ Workplace ❑ Store
❑ Your Summer Residence ❑ Hotel/Motel ❑ Recreation Area
❑ Other (please specify) ______

3 Street Address ______
Nearest Intersection/Landmark ______
Town ______ State/Province ______ Zip ______

4 ❑ Your Primary Residence ❑ Workplace ❑ Store
❑ Your Summer Residence ❑ Hotel/Motel ❑ Recreation Area
❑ Other (please specify) ______

5 ❑ I-89 ❑ I-91 ❑ I-93 ❑ Route 4 ❑ Route 7 ❑ Route 9
❑ Other (please specify) ______

6 ❑ Work Commute ❑ Business Related
❑ Shopping ❑ School
❑ Recreation ❑ Other (please specify) ______

7 ❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5 or more

8 ❑ Passenger vehicle/motorcycle ❑ Pick-up truck/van
❑ Truck (2+ axles, more than 4 tires) ❑ Other ______

9 Comments ______

FIGURE 1 Sample postcard used in survey.

motorist who wanted to know details about the surveys, was prepared. The project description sheet provided motorists with the name and telephone number of a contact person within VAOT from whom more information could be obtained.

Survey Station Selection

Site visits were made to identify potential locations along each route where O-D surveys could be conducted. The following criteria were used to identify precise survey stations for each roadway:

- **Sight distance:** The primary consideration in selecting survey locations was safety. This required visibility and sight distance to be evaluated. Survey stations were located on flat, straight stretches of road. Locations near curves, hills, and other obstructions to visibility were avoided.
- **Roadway cross section:** Wherever possible, survey stations were placed where roadway width was at its maximum and in an area with shoulders. This was particularly important for interview sites, where three travel lanes were needed (see section on traffic control plans). In addition, culverts, ditches, utility poles, and private property were avoided. In general, the more room available and the fewer obstacles present, the safer the traffic operations were through the site.
- **Proximity to state line:** Because this survey was aimed at external-to-internal and external-to-external trip data, it was important to be as close as possible to the state line. The greater the distance from the border, the greater was the possibility of intervening land uses that would attract trips that could not be surveyed.
- **Minimizing delays for traffic flow:** Wherever possible, sites were selected at locations where vehicles normally stop. Types of sites that fell into this category were a ferry dock at the origin of one route, signalized intersections, and the U.S. Border Inspection Stations at the Canadian border.

Survey Scheduling

The surveys were conducted on Tuesdays, Wednesdays, and Thursdays during June, for 12 hr at each site (7:00 a.m. to 7:00 p.m.). The following issues were considered in scheduling the surveys:

- **Month:** The choice of month depended on whether typical or peak data were desired. The objective of this study was to obtain data for typical or average conditions. An examination of monthly traffic counts from permanent traffic count stations revealed that traffic conditions in June are generally typical; therefore, June was chosen as the survey month.
- **Day of week:** The choice of day of week also depended on the type of traffic model being developed. The Vermont model is an average weekday daily model. Therefore, surveys were conducted only on weekdays. In addition, Mondays and Fridays were eliminated because of their proximity to the weekend.
- **Time of day:** For safety reasons, it is best to conduct O-D surveys during daylight hours. An analysis of traffic counts from

the permanent count stations revealed that the 12 hr with the most traffic volume were generally between 7:00 a.m. and 7:00 p.m. The permanent count station data indicated that these 12 hr generally accounted for over 75 percent of total daily traffic volume.

Contingency days were included in case specific surveys had to be canceled for any reason. Rain was the primary concern because of its impact on the survey crew and safety and visibility issues. In addition, personnel problems, including absent temporary employees or police, could force cancellation or suspension of operations.

FIELD DATA COLLECTION

Great care was taken to ensure that the surveys were conducted in a safe and efficient manner. Methods for setting up the survey stations, distributing postcards, and conducting interviews, including the safety and training measures taken, are described. In addition, the survey responses are summarized.

Traffic Control Plans

Typical traffic control plans were developed for three different types of sites, as shown in Figures 2 and 3. The three types of sites are two-lane road mail-back postcard site, two-lane road interview site, and Interstate highway mail-back postcard site.

The differences between the plans that are important to note relate to the number of lanes needed for interviews or postcard distribution. On two-lane roads, postcards were distributed to motorists in their normal travel lane, with a small buffer area between the opposing lanes of traffic in which crew members could work. At interview sites, three lanes were created using traffic cones, two in the direction of the survey and one in the opposite

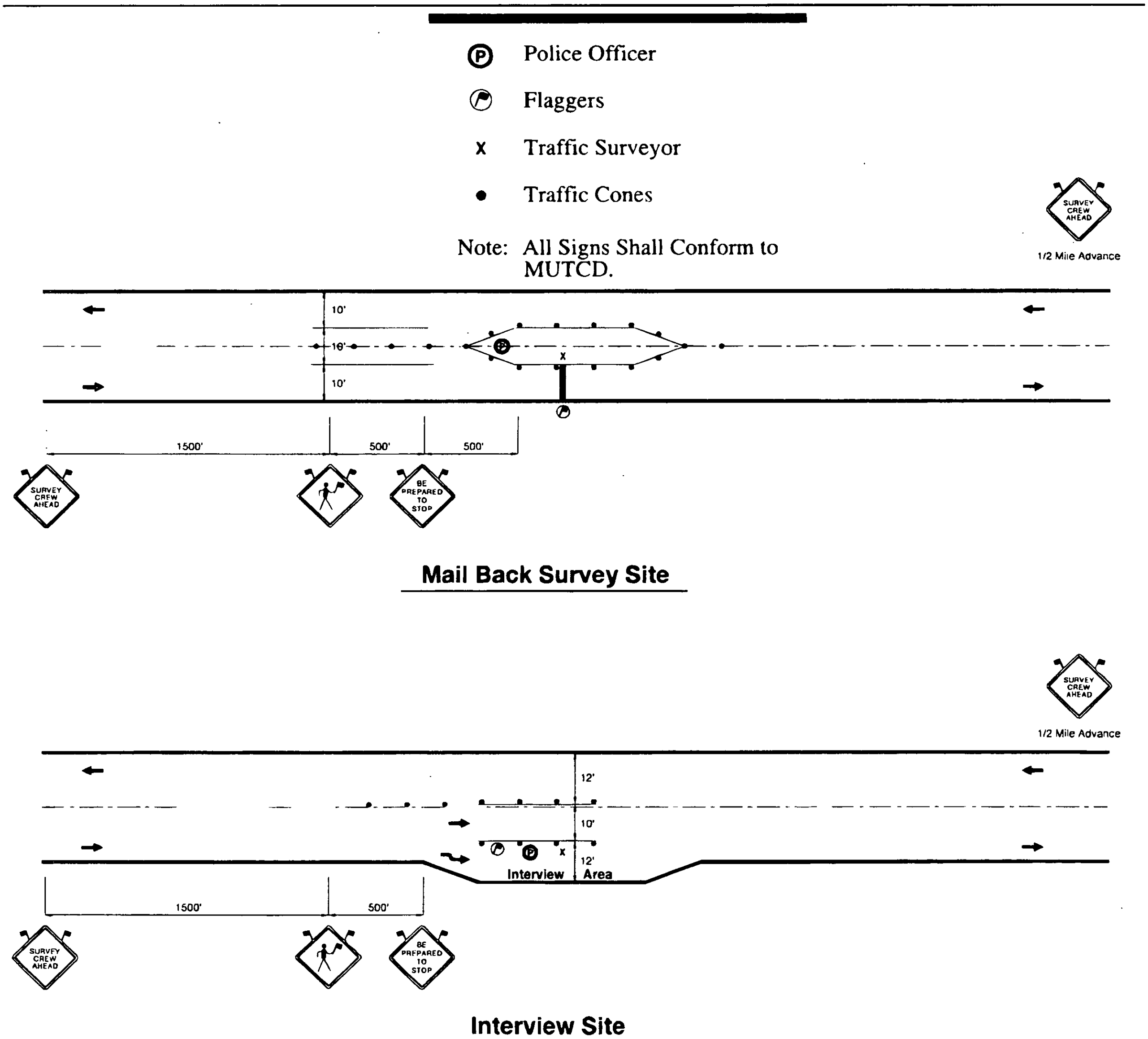


FIGURE 2 Two-lane-roadway O-D survey sign package.

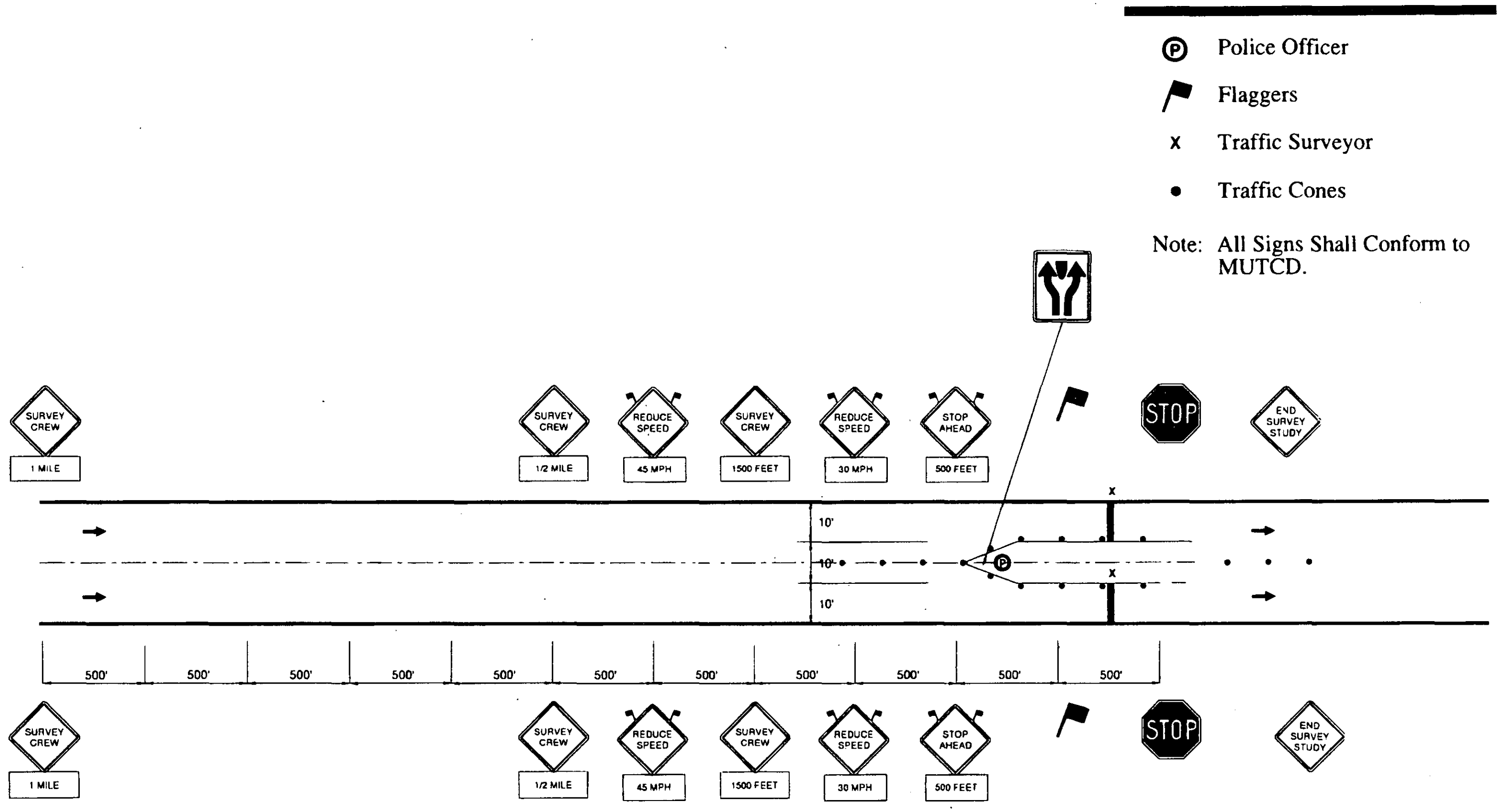


FIGURE 3 Interstate O-D survey sign package.

direction. The rightmost lane in the survey direction was the designated interview lane. The center lane was used as a bypass lane for motorists who were not stopped. The third lane was used by vehicles traveling in the opposite direction to the survey. The Interstate site was a mail-back postcard site. Cards were distributed to vehicles in both lanes rather than in just one in order to minimize delay and to eliminate lane distribution bias in the sample.

All traffic signs used for the surveys conformed to the standards set in the *Manual on Uniform Traffic Control Devices* (*5*). The signs included advance warning signs about the survey crew and the flagger and a sign instructing motorists to be prepared to stop. The first sign was placed approximately 0.5 mi in advance of the survey station. The Interstate sign package was more extensive, with the first sign placed 1 mi from the site. The signs instructed motorists to reduce speeds and to come to a complete stop at the survey station. Other on-site safety measures included

- Orange reflective safety vests for all crew members,
- Trained flaggers to direct traffic through the survey site,
- Presence of police officers and their vehicles at all sites where vehicles were stopped by crew members for the survey, and
- Avoidance of the roadway by crew members on break.

Postcard Distribution

The number of crew members required to distribute postcards ranged from one to four, depending on the traffic volume. The one exception was the Interstate site, which had eight people distributing cards (four in each lane) in order to minimize backups. The crew members were spread out through the survey station to allow cards to be distributed to more than one vehicle at a time. When a crew member was available, the next vehicle would be directed into the survey station by the flagger. The postcards were bundled in envelopes according to the hour in which they were to be distributed. The site supervisor was responsible for switching envelopes between hours.

Interviews

The number of people required to conduct interviews ranged from two to three, depending on the traffic volume. The interviewers were spread out through the survey station to allow more than one motorist to be interviewed at a time. When an interviewer was available, the next vehicle would be directed into the station by the flagger. The site supervisor was responsible for checking interview response sheets throughout the day to make sure that the interviewers were recording the answers correctly.

Summary of Surveys

A log of the number of postcards completed and returned, or the number of interviews conducted, was maintained by time of day for each site. The responses were cleaned up and errors were corrected whenever possible. Usable responses were separated from those that could not be interpreted. The review indicated that over 90 percent of all responses received (postcards and interviews) were usable. The most common types of errors observed included

- Answers recorded in the wrong place,
- Incomplete street addresses and ZIP codes, and
- Information provided for the round trip rather than the one-way trip as requested (although this occurred on less than 5 percent of all returned postcards).

The summary of survey responses is presented in Table 2. The number of usable responses was compared with the observed traffic volume at each site. Traffic counts were conducted at all sites for 72 hr around the survey period. It should be noted that the traffic volumes used to estimate sample size represented the AADT for recent previous years and are different from the traffic volumes presented in Table 2, which represent the actual traffic volumes on the survey day. Overall, approximately 24 percent of all the cards distributed were returned and found to be usable. The returns for each site varied from 9 to 33 percent. Over 95 percent of all the interviews conducted were usable. The comparison of the number of surveys with the 12-hr traffic data indicated that the usable surveys included just under 20 percent of the total traffic volume.

Generally, the lowest response rate, ranging from 9 to 20 percent, was observed for the Canadian border sites. The average for all Canadian border sites was less than 14 percent. The reasons for the lower response rate of these sites may have been the need for additional postage on postcards mailed outside the United States and the language difference.

Minimizing Data Biases

Survey methods are generally susceptible to biases. Measures were taken to minimize biases during the planning and implementation phases, including the following:

- Vehicles were selected at random, including trucks;
- On four-lane roadways, vehicles were surveyed in both lanes or traffic was directed into one lane before being surveyed;
- Cards in two languages were prepared to avoid nonresponse due to language difference; and
- Sites were selected to represent different functional classifications, traffic volumes, and geographical areas.

In addition, biases are also introduced by the unwillingness of certain subgroups within the population to respond to these types of surveys. Nonresponse bias is likely to be higher for the mailback surveys than for the interview surveys. The impact of this type of bias will be analyzed during the analysis phase of this study.

STUDY COSTS

A summary of estimated costs for design and implementation of the O-D survey is presented in Table 3, which shows that the total study cost was approximately $121,500. The cost includes consultant staff and VAOT employee labor, temporary labor, traffic control expenses, printing, postage, and expenses for food, travel, and lodging. Generally, one VAOT employee was present at each site that had volumes over 2,000 AADT. The cost estimate does not include expenses incurred by VAOT staff, as those expenses are not known. The costs also do not include data entry and extensive data analysis, which are currently under way.

The costs were broken down by task and by survey type for comparison, as shown in Table 3. The labor costs for the design/planning task accounted for the largest percentage of effort, represent-

TABLE 2 Summary of Results of O-D Survey

Route Number	Bordering State	One-Way 24 Hour Traffic Volume*	One-Way 12 Hour Traffic Volume*	Number of people Interviewed	Usable Interviews	Response Rate**	Number of Post Cards Distributed	Number Returned	Usable Cards	Response Rate***	Usable Sample As Percent of 12 hour Traffic Volume
Interview Sites											
US 2	NY	2,429	1,822	511	503	98.4%	-	-	-	-	27.6%
VT 314	NY	936	754	308	307	99.7%	-	-	-	-	40.7%
VT 17	NY	883	514	456	443	97.1%	-	-	-	-	86.2%
VT 149	NY	2,135	1,626	630	623	98.9%	-	-	-	-	38.3%
VT 114	NH	1,487	1,076	612	600	98.0%	-	-	-	-	55.8%
US 2	NH	1,828	1,412	747	699	93.6%	-	-	-	-	49.5%
VT 123	NH	2,482	1,954	676	563	83.3%	-	-	-	-	28.8%
VT 8/100	MA	1,477	1,134	354	349	98.6%	-	-	-	-	30.8%
VT 142	MA	609	473	228	220	96.5%	-	-	-	-	46.5%
Mail-back Sites											
US 4	NY	3,229	2,309	-	-	-	1,941	456	420	21.6%	18.2%
VT 9	NY	4,871	3,742	-	-	-	2,815	735	678	24.1%	18.1%
VT 10A (EB)	NH	9,085	7,307	-	-	-	2,556	913	847	33.1%	11.6%
VT 10A (WB)	NH	9,090	7,210	-	-	-	2,515	850	755	30.0%	10.5%
NH 119	NH	5,423	4,191	-	-	-	2,241	596	526	23.5%	12.6%
NH 25	NH	1,161	941	-	-	-	769	194	167	21.7%	17.7%
US 7	MA	4,276	3,385	-	-	-	1,029	175	142	13.8%	4.2%
I-91	MA	7,236	5,427	-	-	-	4,304	1,067	1,000	23.2%	18.4%
I-89	CAN	1,696	1,145	-	-	-	768	110	96	12.5%	8.4%
VT 139	CAN	257	243	-	-	-	243	54	48	19.8%	19.8%
I-91	CAN	756	620	-	-	-	595	108	92	15.5%	14.8%
VT 147	CAN	454	376	-	-	-	338	42	30	8.9%	8.0%
Totals		61,800	47,661	4,522	4,307	95.2%	20,114	5,300	4,801	23.9%	19.1%

*Traffic counts conducted in June, 1994 while surveys were being conducted.
**Number of usable interviews divided by the total number of interviews.
***Number of usable postcards divided by the total number of postcards distributed.

TABLE 3 Cost Summary of O-D Survey

		Mail-back Postcards	Interviews	Total
Labor				
	Design/Planning and Site Visits	31,100.00	25,900.00	57,000.00
	Data Collection	21,700.00	18,800.00	40,500.00
	Subtotal	52,800.00	44,700.00	97,500.00
Expenses				
	Travel, food, misc.	5,200.00	4,300.00	9,500.00
	Traffic Control Equipment	1,800.00	1,500.00	3,300.00
	Police Details	2,400.00	2,700.00	5,100.00
	Printing, postage	6,000.00	100.00	6,100.00
	Subtotal	15,400.00	8,600.00	24,000.00
Total Cost		$68,200.00	$53,300.00	$121,500.00
Number of Usable Responses		4,800	4,300	9,100
Cost Per Usable Response		$14.21	$12.40	$13.35

The split between mail-back and interview techniques was based on exact hours in all cases where they could be determined. In all other cases, the split was estimated proportionally.

ing approximately 47 percent of the project cost. The labor cost for conducting the survey was approximately 33 percent, and expenses accounted for the remaining 20 percent. The labor cost for the data collection task was lower than design and planning labor cost because temporary agency employees working on an hourly basis were used to conduct the surveys.

The breakdown between mail-back postcard surveys and interview surveys was included to illustrate the difference in cost between the two techniques. The number of usable surveys included the mail-back postcards that were filled out correctly by the respondents and the interviews that were completed correctly. The cost per usable response, reflecting the costs incurred in order to get one usable postcard or one usable interview, was estimated to be $13.35. The estimated cost per usable response for the mail-back postcard survey was $14.21, compared with $12.40 for the interview survey.

STUDY FINDINGS

One of the most interesting parts of this study was the opportunity to compare the costs of the different survey techniques involved. Table 3 indicates that the cost per usable response for the mail-back postcard survey ($14.21) is greater than that for the interview survey ($12.40). However, much of this difference results from the fact that in this study interviews were conducted on low-volume roadways (less than 5,000 vehicles per day [vpd]). On the other hand, mail-back postcard surveys were conducted for all higher-volume roads (greater than 5,000 vpd) and lower-volume roads where interviews could not be conducted. As indicated earlier, the response rate was lowest for Canadian border sites where mail-back postcard surveys were conducted, which contributed to higher overall cost per usable postcard survey. The ferry site (VT 314), which was an interview site, did not require any traffic control because the motorists were interviewed while they were waiting in the holding area before boarding the ferry. Fewer interviewers were required because more time was available for conducting interviews, further contributing to the lower costs for interview surveys.

It was considered more appropriate to compare costs for sites with similar traffic volumes and conditions. The first group of sites included traffic volumes from 2,000 to 3,000 vpd, and the second group included 4,500 to 6,500 vpd. The cost comparisons are presented in Table 4. Although data were not available to compare costs for higher-volume roads (over 6,500 vpd), mail-back postcard cost per usable response was estimated for a representative higher-volume road. This cost estimate is also presented in Table 4.

It is generally more cost-effective to conduct interview surveys on lower-volume roads (5,000 vpd or less). The comparison indicated that the cost per usable response for mail-back postcard surveys decreased as traffic volumes increased. Cost per usable response for higher-volume roads like I-91 were as low as $8.50. This Interstate site required the largest number of crew members and police officers, and the traffic control plan was the most elaborate, resulting in a higher total cost for the site. However, the cost per usable response was much lower, indicating that the cost per usable mailback survey drops as traffic volume increases.

Additional study findings include the following:

- Interview sites generally required more traffic control, more training time, and more supervision of crew members than mail-back postcard sites;
- Interviews could generally be conducted in less than 1 min;
- Police assistance is desirable for all sites where traffic is stopped on the road for the survey;
- The mail-back postcard technique is generally not appropriate for lower-volume roads because the lower response rate may affect sample size requirements;
- Interview surveys were generally cost-effective for lower-volume (5,000 AADT or less) roadways.

TABLE 4 Comparison of Interview and Postcard Techniques for Selected Roadways in Vermont

Traffic Volume Range: 2,000-3,000 Vehicles per Day	Interview (VT 114)	Mail-back Postcard (NH 25)
Two-way Traffic Volume	2,974	2,322
One-way Traffic Volume	1,487	1,161
Usable Responses	600	167
Interviews Conducted or Postcards Distributed	612	769
Response Rate	98.0%	21.7%
Estimated Cost per Site	$5,600.00	$5,300.00
Estimated Cost per Usable Response	$9.35	$31.75
Traffic Volume Range: 4,500-6,500 Vehicles per Day	**Interview (US 2 NY)**	**Mail-back Postcard (US 4 NY)**
Two-way Traffic Volume	4,860	6,450
One-way Traffic Volume	2,430	3,230
Usable Responses	503	420
Interviews Conducted or Postcards Distributed	511	1,941
Response Rate	98.4%	21.6%
Estimated Cost per Site	$5,600.00	$6,150.00
Estimated Cost per Usable Response	$11.15	$14.65
Traffic Volume Range: 14,500 Vehicles per Day	**Interview**	**Mail-back Postcard (I-91 MA)**
Two-way Traffic Volume	N/A	14,500
One-way Traffic Volume	N/A	7,250
Usable Responses	N/A	1,000
Interviews Conducted or Postcards Distributed	N/A	4,304
Response Rate	N/A	23.2%
Estimated Cost per Site	N/A	$8,500.00
Estimated Cost per Usable Response	N/A	$8.50

The study provided an opportunity to compare two popular O-D data collection techniques—mail-back surveys and interview surveys. Each method has advantages and disadvantages and cost concerns. Transportation planners should evaluate study requirements and site restraints before selecting a survey method to suit their needs. In general, this study concludes that the interview technique is more cost-effective for lower-volume (less than 5,000 vpd) roadways.

ACKNOWLEDGMENTS

The authors would like to thank the Vermont Agency of Transportation for the opportunity to conduct the study. In particular, the assistance provided by David Scott and the VAOT project team was extremely helpful. Contributions by Charles C. Crevo, the VHB project manager for this project, were of great value.

REFERENCES

1 *Manual of Transportation Engineering Studies*. Institute of Transportation Engineers, Washington, D.C., 1994.
2. Miller, K. S., et al. Using Video Technology to Conduct 1991 Boston Regional External Cordon Survey. In *Transportation Research Record 1412*, TRB, National Research Council, Washington, D.C., 1992.
3. Stokes, R. W., and T. Chira-Chivala. Design and Implementation of Intercity Origin-Destination Surveys. In *Transportation Research Record 1236*, TRB, National Research Council, Washington, D.C., 1989.
4. Hajek, J. J. *Optimal Sample Size of Roadside Interview O-D Surveys*. Ontario Ministry of Transportation and Communication, Ontario, Canada, Jan. 1977.
5. FHWA. *Manual on Uniform Traffic Control Devices*. U.S. Department of Transportation, 1988.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Design and Conduct of a Statewide Household Travel Survey in Vermont

CHARLES C. CREVO, RAYMOND S. NIEDOWSKI, AND DAVID J. SCOTT

In their efforts to develop a statewide travel demand model, the Vermont Agency of Transportation and its consultant, Vanasse Hangen Brustlin Inc., rejected the option of using data gathered from another area because of the unique character of Vermont and the lifestyle of its residents. This choice was further supported by the need to develop regional submodels within the statewide structure. The agency decided to conduct travel surveys, one of which would include a representative sample of households in the state. A discussion is presented of the process followed in collecting the data required to develop the models. Because of the renewed interest in travel demand modeling created by the Intermodal Surface Transportation Efficiency Act of 1991 and the Clean Air Act Amendments of 1990, practitioners should find the activities and costs associated with this effort helpful in making decisions regarding household socioeconomic and travel inventories they might be planning to undertake. The paper focuses on sample selection, form design, printing, mailing, receipt, and quality control associated with the administration of the survey.

The information presented in this paper regarding techniques, results, and costs of a statewide mail-out/mail-back household travel survey conducted in 1994 should be beneficial to practitioners considering, or about to undertake, some form of household-level travel data collection effort. This paper presents a discussion of the process followed in collecting data required to develop a statewide travel demand model in Vermont and the cost of the activities, which include sample selection, form design, printing, mailing, receipt, and quality control associated with the administration of the surveys.

HISTORY

Since the early days of travel demand model development, household surveys have been the major source of socioeconomic and travel-related data. The information gathered through inventories of household and trip-making characteristics has been used to establish three basic interrelationships of the traditional four-step modeling process, namely, trip generation and distribution and mode share.

In the early 1960s, large-scale surveys were conducted to gather the information needed to develop travel demand models. Initial data collection efforts used the personal home interview technique, and the surveys were conducted in a door-to-door campaign, a method that was time consuming and costly. In the late 1960s and early 1970s, other techniques were applied to reduce the costs associated with these data compilation efforts. Popular alternative approaches to obtaining household characteristics and travel data were telephone and mail surveys. Each had advantages and disadvantages, although the deciding factor was usually a combination of data reliability and cost.

In 1991 the passage of the landmark Intermodal Surface Transportation Efficiency Act stimulated a renewed interest in travel demand modeling and in obtaining current socioeconomic and trip-making relationships, particularly with regard to time-of-day travel and information required to develop disaggregate models and other innovative applications. Peak-hour estimating capabilities are particularly important to satisfy some of the requirements of another significant piece of legislation, the Clean Air Act Amendments of 1990.

BACKGROUND

In their efforts to develop a statewide travel demand model, the Vermont Agency of Transportation (VAOT) and its consultant, Vanasse Hangen Brustlin Inc., (VHB), decided that it would be necessary to conduct travel surveys, one of which would include a representative sample of households in the state. The VAOT chose not to use data from other studies because of the statewide-regional nature of the model structure and because of the unique character of Vermont and the lifestyle of its residents. The Nationwide Personal Transportation Survey was also considered, but the small sample size in Vermont would require use of regional data, which include information from highly urbanized areas in New England.

In recent years, a number of data collection techniques have been developed and tested. Some of the more advanced are diary surveys and panel surveys. These data collection efforts can be for a point in time or an extended period. Methods to gather the information rely on personal, mail, or telephone contact with various follow-up techniques to enhance return rates. A number of documents and publications were reviewed, and various techniques for conducting household surveys were evaluated. The mail-out/mail-back method was selected for the following reasons:

- Personal interviews are time consuming and costly.
- Telephone interviews, although efficient and requiring fewer household contacts to obtain the required sample, generally use a random dialing technique, which does not guarantee a geographically representative sample. Also, competition by myriad telemarketing efforts of commercial enterprises could discourage participation by some households.
- The mail technique, although requiring a mailing to a greater number of households to obtain the necessary sample size, would ensure the representative samples needed for model development. However, one drawback of the mail-out/mail-back survey tech-

C. C. Crevo and R. S. Niedowski, Vanasse Hangen Brustlin, Inc. 101 Walnut Street, Watertown, Mass. 02172. D. J. Scott, Vermont Agency of Transportation, 133 State Street, Montpelier, Vt. 05602.

nique is that nonresponse bias cannot be evaluated because what is missing is unknown.

SAMPLE SELECTION

One of the objectives of the sample selection was to obtain a geographical representation. A target of 1 percent of the dwellings was established to provide sufficient data to develop a cross-classification approach to trip generation, trip length characteristics for trip distribution model parameters, and mode share characteristics. According to the 1990 Census Transportation Planning Package, there are approximately 218,900 dwellings in Vermont. Assuming a 10 to 15 percent response rate, approximately 28,200, or 13 percent, were selected to receive questionnaires.

The next task was to locate a list of mailing addresses from which a sample could be selected. Candidate sources were telephone directories, tax assessor maps, and commercial mail houses that specialize in such lists. The last proved to be the best choice because this service included generation of mailing labels and a control list. The company selected for this task also printed the survey questionnaire, inserted serial numbers on the mailing labels, and assumed responsibility for mailing the forms. The printer was provided a list of towns in Vermont with the number of deliverable addresses required in each.

FORM DESIGN

Two basic types of household information are necessary for model development: demographic characteristics of members of the specified households and trip-making characteristics of household members. The first includes information on age, sex, employment status, number of automobiles available, and income level. These data are required to establish trip rates for various socioeconomic parameters. Such rates can then be used by the state and regions to forecast travel demands for specific geographic areas. Required trip information includes origin and destination data by time and place, travel mode, vehicle occupancy, and trip purpose. Design of the survey form would have to incorporate each of these needs.

A preliminary survey package was developed that consisted of the following components:

- A letter from the Governor of Vermont soliciting cooperation and explaining the purpose of and need for the survey,
- A set of instructions on how to complete the survey package,
- A form soliciting demographic information about the household and its members,
- A form with questions on transportation prepared by the VAOT plus space for the respondent to express an opinion on anything related to transportation,
- A series of forms asking for information on trips made by household members on one particular survey day, and
- Maps to be developed for use by the respondent to identify trip origins and destinations by traffic analysis zone.

The initial concept was to develop the survey package as an unbound series of individual sheets that would be used to solicit trip information from household members aged 5 or more, to ask for such information for the day before receipt of the survey package in the mail, and to use the map technique to minimize subsequent effort required to translate address or place-name information into the traffic analysis zone format required by the model. Other aspects of the survey identified during this early phase included

- Exploration of the possibility and desirability of an incentive (monetary prize or premium) to help increase the survey response rate;
- Provision of a toll-free telephone number so that potential respondents could contact the survey team for assistance, additional forms, or other reasons;
- Use of a separate diary form for each household member;
- Use of informal language rather than the more formal and technical verbiage; and
- Use of a question-and-answer format for conveying the required instructions.

Several meetings were held to solicit comments on early design efforts and refine the survey procedures. This process resulted in the following decisions:

- To use a booklet-type format that could be mailed directly, eliminating the need for separate envelopes. This approach reduced the survey's complexity and cost, as well as the probability that individual sheets would be lost or damaged by respondents.
- To use an 8½-in. by 11-in. format for readability and to reduce the mailed booklet's size by folding it to 8½ in. by 5½ in. to reduce postage costs. Reversing the fold and resealing by the respondent after completion of the survey resulted in a returnable package.
- To limit the booklet's composition to four 11-in. by 17-in. sheets (16 printed pages of 8½ in. by 11 in.) to minimize weight and reduce postage costs.
- To avoid mention of VHB in the survey package to emphasize that this was a VAOT survey. It was decided that a better response would be achieved with a return address to VAOT in Montpelier rather than to VHB in Massachusetts. VAOT forwarded all returns to VHB for processing.
- To minimize printing costs by using only one color and one paper weight for all pages of the survey package so that only one press run would be required.
- To abandon the concept of the traffic zone map because of time constraints and to keep things as simple as possible for the respondent. Because of the statewide nature of the survey, only partial map coverage was practical, necessitating a combined zone/address technique that could have caused confusion. Also, map reading and interpretation could have been a problem for some respondents.
- To abandon the incentive approach because cost considerations of a large mailout (more than 28,000 surveys) precluded giving everyone a premium and because the likely benefit derived from distributing a few small prizes was thought to be limited.
- To provide a toll-free telephone number with instructions to call during normal business hours. An answering machine was activated after working hours to provide continuous access.
- To use informal language and a question-and-answer format.
- To retain the separate trip diary form technique and provide extra forms for large households or for household members making many trips.
- To designate a nonspecific survey day rather than a predetermined day for each household because of the relative unpredictability of the delivery schedule of third-class mail, which, according to the post office, is typically 1 to 2 weeks.

- To designate the survey day for each household as the day immediately after receipt of the survey rather than before. This reduced recall problems associated with recording trips made the previous day.
- To solicit only weekday information because the model would have to replicate typical daily travel, which in most cases and geographic areas includes weekday commuter travel.
- To stagger mailings to obtain trip information for a sampling of weekdays rather than for one particular day.
- To solicit information only for "reportable trips," which are important to the development of the vehicular and transit trip models. Incidental trips, duplicate trips, and trips typically included only for the sake of comprehensiveness were eliminated to the extent possible to reduce the amount of information the respondent had to record.
- To eliminate mention of an age cutoff for trip making. All trips by members of a household were considered important regardless of the person's age.
- To provide a sample form to demonstrate how to complete the trip diary.
- To exclude college students from the survey. This decision was based on considerations of time (most college students would be either finished with school or focused on final examinations by the time of the survey), effort (they would require special technical and administrative procedures to handle), and size of group (they made up only a very small part of the total statewide universe).

SURVEY METHOD

VHB decided that a staggered mail-out system was desirable to produce trip information on each weekday. The mail-out schedule established with the vendor was to deliver survey booklets to the post office over several days. Because Wednesday, May 11, was the first possible day for mailing, Friday, May 13, was specified as the last mail-out day. This schedule allowed sufficient time between receipt of the survey and the Memorial Day weekend, because the post office was estimating a 1-week to 10-day delivery time frame rather than up to 2 weeks, as mentioned earlier. Although it was not possible to control the delivery schedule, VHB estimated there would be a sufficient natural stagger across the state to allow for survey responses representative of each weekday. This mail-out schedule was accomplished by the selected print/mail house as planned. Approximately 4,700 survey forms were sent out on Wednesday, May 11; 12,500 on Thursday, May 12; and 11,000 on Friday, May 13. Distribution by town and region followed the calculated sample-size targets as closely as possible, although some modification was necessary because of recent ZIP code reassignments made by Vermont postal authorities.

On Monday, May 16, VHB began to receive telephone calls on the toll-free number established for the survey. Additional calls came in over the next week or so, indicating that most people received the forms within a few days of mail-out rather than the 2-week worst-case time frame initially estimated by the post office.

By June 1, 46 calls had been received. The nature of the calls was mixed. Many were to request clarification about the definition of a trip, to say they made few trips or none on the survey day, or to ask what to do if they missed the survey day. Some requested guidance because the form was sent to a deceased, ill, or elderly relative. Several had concerns about privacy and said they would not fill out the form or would remove the label before sending it in. A few had questions or comments about a gasoline tax question. A small number said they had received more than one form or received the form at their place of business. Two thought the survey was a waste of time and money and refused to complete the forms.

As surveys arrived at VAOT, they were packaged and sent to VHB in batches, where every returned survey form was logged by staff members who recorded the serial number, town name, and ZIP code from the form's mailing label. A custom computer spreadsheet was developed for this purpose, which facilitated sorting the form's contents; the spreadsheet also automatically kept a running count of the number of entries. In the event that the label was removed by a respondent for privacy reasons, the inside of the form was inspected for town and ZIP code information and, if available, recorded. In this case a notation was made in the spreadsheet indicating that usable information was available despite the lack of a mailing label. It is ironic that most respondents who removed the mailing label, presumably for privacy reasons, did not hesitate to provide their telephone number, which was requested to allow a follow-up call if necessary. If the respondent provided no information for whatever reason, the form was logged as a return but recorded as having no usable information. To prevent double counting, each form was marked to indicate that it had been processed. The rate of return of survey booklets by time after mailing is shown in Figure 1.

A small percentage of forms were still being received more than a month after they were mailed. On these late returns, some respondents apparently attempted to provide trip information for the day after they received the survey, as requested. Others ignored that instruction and completed the form for a more recent day, presumably because it was easier to remember. Some respondents filled in only household-related information, perhaps thinking that too much time had passed since they received the booklet to remember past trips but not wanting to ignore the timing instruction.

RESULTS

ZIP codes from the above-mentioned tabulation were copied to another spreadsheet, which was developed to summarize return statistics by individual towns within each Regional Planning Commission (RPC) area, as well as for the entire state. The table included, for each town/ZIP code, the original target sample desired, the actual number of survey forms mailed and returned, and the percent returned. In some cases towns were combined with others when the local ZIP code system dictated such an arrangement.

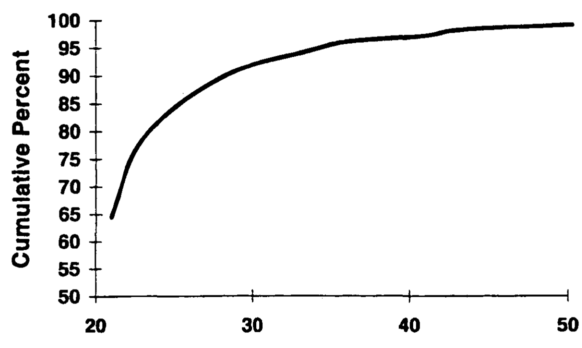


FIGURE 1 Survey returns by time.

The spreadsheet was set up to automatically update each new batch of ZIP codes and modify the return statistics by location. Table 1 summarizes the key information by RPC.

The 2,447 returned survey forms represent 1.1 percent of the total number of households in Vermont. On a regional basis, the Addison RPC had a return rate of 13.7 percent, whereas response rates for the other RPCs were generally in the 7 to 9 percent range.

The only presurvey activity was a news release by the VAOT public relations unit to announce the survey. Given the complexity of the information requested and the lack of an initial contact phase to identify likely respondents or solicit cooperation, the returns were sufficient for the purposes of this project. No follow-up telephone calls or supplementary mailings were made to increase the response rate. A meaningful increase in returns would be forthcoming only through significant additional expense.

The cost to design and conduct the household survey was approximately $17.50 per returned survey. The cost breakdown is shown below. The first two components are best estimates, and the last two represent actual costs.

Item	*Cost($)*
Planning and design of survey form and presurvey arrangements	10.70
Production (develop mailing list, print, bind, etc.)	3.90
Third-class outgoing postage	2.30
First-class return postage	0.60
Total	17.50

Of the forms returned, approximately 79 percent were fully usable and an additional 17 percent were made usable through an editing process. The remaining 4 percent were unusable.

CONCLUSIONS

The results of the Vermont statewide household survey indicate that the mail-out/mail-back technique for obtaining information on household demographic and travel characteristics without advance contact with potential respondents represents a viable and cost-effective approach. It is a relatively simple procedure to organize and administer, and it results in data with an acceptable level of reliability. However, given the lack of advance contact and the complexity associated with this type of survey, the practitioner can expect a return rate of 8 to 10 percent compared with the typical 15 to 25 percent associated with less complex mail-type surveys. VHB was prepared for a lower-than-average return because of some small-scale advance testing. Advance contact by mail or telephone could have improved the return rate, although at a higher cost. Also, an incentive program of some sort could be beneficial in cases for which the survey area is of reasonable size. This household survey involved the entire state of Vermont, and the cost to implement a meaningful incentive program was prohibitive. Follow-up contact by mail or telephone could also improve the return rate, again at a higher cost. The survey practitioner must evaluate project needs and determine whether achieving a higher response rate is worth the additional cost.

For the Vermont survey, the selected format worked well. Everything required was contained in one package, precluding the need for separate mailing envelopes and upfront handling. The 8½-in. by 11-in. size was large enough for readability without sacrificing the postage benefits of a low-weight piece. The booklets held up well for the most part and, except where purposely detached by some respondents for privacy reasons, had no missing pages on return. Based on a preliminary review of the returns, fewer household member trip diary pages would have been sufficient than the eight provided.

Caution is advised in the development of a mailing list. The selected method of using a printing vendor to provide an address list resulted in inconsistencies. For example, the address list was approximately one-half the total number of households in the state. The list was supposed to be a "clean" one, having only "deliverable" addresses. If this approach is used, one should be certain that the smaller list has a geographic distribution similar to the region to be surveyed thereby avoiding a potential geographic bias, which was the purpose of this project. A second condition required the shifting of addresses among towns because of the lack of sufficient

TABLE 1 Survey Return Statistics

Regional Planning Commission	Households	Mailings	Percent	Returns[1]	Percent
Addison	13,606	1,714	12.60	234	13.7
Bennington	16,461	2,095	12.73	173	8.3
Central Vermont	27,264	4,248	15.58	330	7.8
Franklin - Grand Isle	21,385	2,555	11.95	179	7.0
Lamoille	9,872	1,441	14.60	108	7.5
Northeastern Vermont	30,849	3,593	11.65	332	9.2
Rutland	30,646	3,833	12.51	306	8.0
Southern Windsor	13,675	1,988	14.54	175	8.8
Two Rivers - Upper Valley	26,923	3,144	11.68	285	9.1
Windham	28,312	3,584	12.66	303	8.5
TOTALS	**218,993**	**28,195**	**12.87**	**2,425**	**8.6**

1. Unknown zip codes account for an additional 22 returns not assignable to an RPC addresses in certain communities to meet the desired sample size. This was acceptable for the Vermont survey because it was more important to achieve a representative sample at the regional level than by individual town. A third inconsistency was that peculiarities in the state's postal system necessitated combining some ZIP codes. In one case, substitute forms had to be sent to one town that had been divided into five ZIP codes. Finally, some forms were returned with notes indicating that the addressee had been deceased for some time, raising a question regarding the age of the list. The advantage of using a vendor to provide an address list is that they typically rent the list for one-time use from a commercial provider and are responsible for dealing with such anomalies. However, if the vendor is less reliable or conscientious than the one selected by VHB, one might never learn about these issues. A thorough check of vendor credentials, along with a detailed investigation of its mailing list, particularly with regard to the issues identified above, is essential at the selection stage.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Traffic Data Collection Using Video-Based Systems

JAMES A. BONNESON AND JOEL W. FITTS

Sources of bias and variability in the measurement of time-based traffic events by using video recording-and-playback systems are described. Bias can affect accuracy, whereas the lack of precision can increase the minimum sample size. Equations are described that can be used to estimate the adjustment needed to remove bias from the recorded data and to estimate the standard deviation of the measurement process. The equations are sensitive to the method of extracting the individual traffic event times from the videotape. These methods include manual extraction, which uses a frame-by-frame analysis, and automated extraction, which uses a video imaging system to analyze the tape during playback at normal speed. The manual method is found to yield less bias and lower variability than the automated method.

One of the principal elements in the design of a traffic data collection plan is event measurement accuracy and precision. The need for accuracy is obvious: measurements must be accurate to be useful for either model calibration or system evaluation. Threats to accuracy typically stem from some distortion in the measurement process that introduces a bias in the measured quantity relative to its true value. This bias can stem from a variety of sources (e.g., measurement device out of calibration, measurements taken in the wrong plane of reference) and typically can be eliminated by device calibration in advance of the study or by adjustment of the data after the study.

Measurement precision is also important to study design. The precision of an estimate of the population mean is dependent on the number of observations, the variance in the population, and the variance of the measurement process. The relationship among these variables is

$$N = \left(\frac{t_{\alpha/2}\, s_p}{e} \right)^2 \quad (1)$$

where

- N = sample size needed for a $1 - \alpha$ level of confidence,
- $t_{\alpha/2}$ = test statistic corresponding to $\alpha/2$ (two-tail test) for normal distribution,
- s_p = pooled standard deviation, and
- e = permissible error (or precision) of the estimate of the true population mean.

The pooled standard deviation combines the variance in the population and the variance in the measurement process as

$$s_p = \sqrt{s^2 + s_m^2} \quad (2)$$

where s is the population standard deviation and s_m is the standard deviation of the measurement process.

The objective of this paper is to describe the sources of bias and variability in the measurement of traffic events by using video-based recording-and-playback systems. Equations developed to describe these sources can be used to estimate the adjustment needed to remove the bias in the recorded data and to estimate the standard deviation of the measurement process (for sample size estimation). For this paper it is assumed that the traffic events of interest have a time base (e.g., headway, speed), that these events have been previously recorded on videotape, and that the data are extracted while the videotape is being replayed.

DATA EXTRACTION CONSIDERATIONS

Data Extraction Methods

Two methods of data extraction are described, manual and automated. Manual extraction entails a frame-by-frame analysis in which event time is computed either by reading the in-picture stopwatch image or by counting the number of elapsed frames and multiplying by the known frame time interval. Either variation of this method requires stopping the videotape at each event to record the time. Automated extraction uses video imaging technology to detect events and record their time of occurrence. The time base for this method is an external clock; thus it does not require that the videotape be stopped to record event time. The video imaging system used for this research is the Autoscope 2003 (software version 3.23), as described by Michalopoulos et al. (*1*).

Field of View

Several terms can be used to describe the field of view obtained from the video camera. These terms include landscape or station and approaching, departing, or overhead.

Landscape and Station Views

Two terms that have been used by Doctor and Courage (*2*) to describe the orientation of the camera in a vertical plane are "landscape" and "station." A landscape view includes the horizon in the upper half of the field of view. The camera is directed outward, and the angle between the center of the field of view and a vertical line is generally 45 degrees or more. This view is best suited to the monitoring of a long [say, greater than 200-ft (61-m)] length of roadway. It can be used to obtain traffic counts and estimates of queue length.

Department of Civil Engineering, University of Nebraska–Lincoln, Lincoln, Nebr. 68588-0531.

A station view is used to obtain traffic data similar to those gathered by a permanent traffic recording station. The camera is directed downward with a viewing angle to the vertical of 45 degrees or less. This view is best suited to the monitoring of a short (say, less than 200-ft) section of roadway. It can be used to obtain precise estimates of headway, travel time, and speed.

Approaching, Departing, and Overhead Views

The view of the traffic lanes can also be categorized as approaching, departing, and overhead. The terms "approaching" and "departing" describe the direction of vehicular travel relative to the camera location. With respect to the video display, traffic moves vertically from top to bottom and from bottom to top for the approaching and departing views, respectively. The overhead view describes the situation in which the camera's field of view is perpendicular to the travel direction and traffic moves horizontally from left to right (or right to left) on the display. These three views are shown in Figure 1.

Some views represent a combination of two of these view types. For example, the camera may be oriented such that traffic flows from the upper left corner to the lower right corner of the video display. The resulting view is often considered desirable because it yields the widest coverage of the roadway. This view would represent a combination of the overhead and approaching view types.

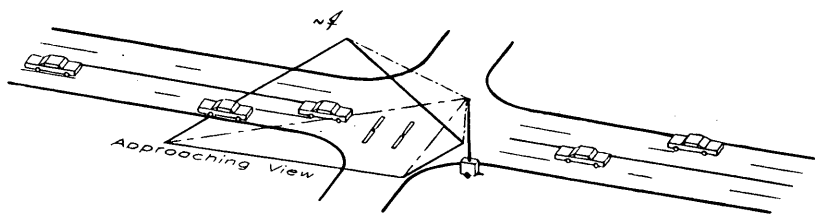
Tape Drag

Times obtained from the automated method or from the manual method (when frame counts are used) must be adjusted to yield the true event time whenever the tape advance mechanism plays more slowly than real time (or drags). Tape drag will also affect event time accuracy when it is measured by an external stopwatch during tape replay (such as when a portable computer is used to record the time that certain keyboard keys are pressed in response to an observed event on the video monitor). The true time can be estimated as

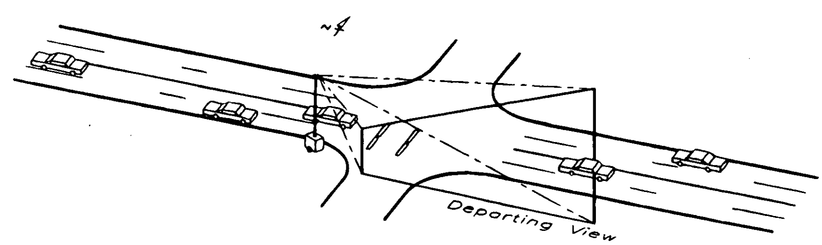
$$T_t = T_r(1 + t_d) \tag{3}$$

with

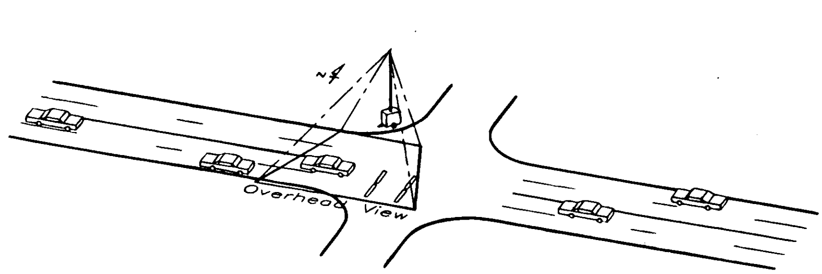
$$t_d = \frac{T_t - T_r}{T_r} \tag{4}$$



a) Approaching view.



b) Departing view.



c) Overhead view.

FIGURE 1 Camera field-of-view categories.

where

T_t = true event time (sec),
T_r = event time measured with an external time clock during tape playback (sec), and
t_d = tape drag adjustment factor (sec/sec).

The adjustment factor t_d can be computed by first videotaping the event and a clock image together (by positioning the clock in the field of view) and then, during playback, comparing the event time obtained from the external clock T_r with the event time obtained from the clock in the video image T_t. The adjustment factor is then computed by using Equation 4. In general, the adjustment factor is a small negative value, indicating that the tape plays more slowly than real time.

As an alternative, T_t and T_r can be obtained from player-recorders that have digital in-picture time clocks and frame-by-frame playback capability. In this situation the tape is replayed frame by frame, stopping at each event to record the in-picture time and the frame count. The frame counts, converted to clock time, represent T_r, whereas the recorded digital time represents T_t. For the player-recorders used in this study the digital frame counts are reported in hours, minutes, seconds, and frames using an internal conversion based on the ratio of 1 sec to 30 frames (1:30). However, the correct ratio is 1:29.970 for color images. Thus, the reported hours, minutes, and seconds first had to be converted back to total frames and then divided by 29.970 to yield the time T_r.

With both techniques a drag of −0.0016 sec/sec was observed for a consumer-grade VHS videotape recorder and −0.00020 sec/sec was found for a professional-grade Hi8 recorder. The drag value for the consumer-grade recorder increased (from −0.0011 to −0.0022 sec/sec) with time into the tape (i.e., the length of tape from the beginning of the reel), whereas no increase was noted for the Hi8 recorder.

One problem with the frame-count technique is that the precision of most camcorder time clocks is limited to the nearest second. This is generally too imprecise to measure traffic event intervals of 10 sec or less. To improve measurement precision, an in-picture time generator device can be inserted into the video feed cable between the camera and the recorder. This generator can then be used to superimpose a digital time image over the camera's video signal. These generators typically provide a precision to 0.01 sec (which exceeds the precision of a video frame, i.e., ≈1/30 sec).

Relative Reference Linewidth

Measurement of traffic event times by either the automated or the manual method is based on the establishment of a reference line on the video display. For the manual method this line can be physically drawn on the video display screen. For the automated method this line would be represented by a video detection zone. In either case the line is typically oriented perpendicular to the travel direction. For example, the line could be located over the stop line of an intersection approach (i.e., physically drawn over the stop line as it is shown on the video display). In this manner the reference line could be used to measure vehicle headways.

The reference linewidth on the video display is relatively thin and presumably constant (as measured in the direction of travel). For example, the Autoscope typically makes its detectors slightly less than 2.5 mm (0.1 in.) wide on a 305-mm (12-in.) display. However, the projection of this line onto the pavement can be much wider, depending on the angle between the camera and a vertical line. A consequence of this increase is a decrease in the precision of event time measurement. For example, in a station view of an intersection stop line the relative width of the projected reference line can be approximately equal to that of the stop line itself, yielding relatively precise measurements of vehicle crossing time. On the other hand, in a landscape view the relative width can exceed that of the observed vehicle, thereby making it impossible to ascertain the exact time when the vehicle crosses the stop line.

The relative reference line width is shown in Figure 2 for a camera positioned 9.8 m (32 ft) above and approximately 6.1 m (20 ft) offset from the centerline of the roadway. Widths are shown for both 6- and 8-mm (0.2- and 0.3-in.) lenses. The camera angle to the vertical was varied between 40 and 60 degrees to generate the data used in the figure. Trigonometric equations were used to transform the reference line coordinates on the video display to the plane of the roadway pavement. These coordinates were then used to compute the transformed linewidth and the horizontal distance between the camera and the equivalent line in the plane of the roadway.

As Figure 2 illustrates, the reference linewidth increases exponentially with the distance between the camera and the reference line. Longer distances correlate with reference lines (or Autoscope detectors) located higher on the video display. The effect on linewidth is most pronounced for the approaching view and least for

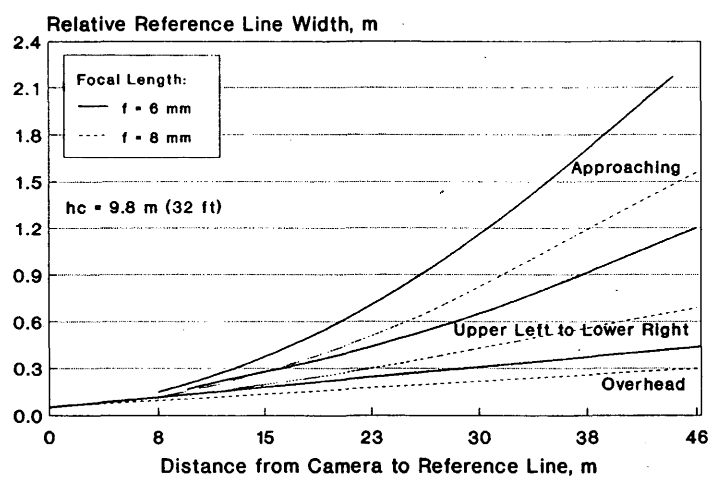


FIGURE 2 Relative reference linewidth (1 m = 3.3 ft).

the overhead view because the distortion of length is greater in the vertical plane of the camera field of view than in the horizontal.

Figure 2 also shows the effect of camera-lens focal length. The more magnified view of the higher-focal-length lens yields shorter linewidths than the lower-focal-length lens. Although this characteristic of higher-focal-length lenses may be advantageous, it must be balanced by the reduced field of view obtained from these lenses. Moreover, the discussion of speed measurement in the next section will suggest that the advantages of a wide field of view may override any benefits obtained from a magnified view. The authors' experience indicates that 6 to 8 mm may be the most useful range of focal lengths, with preference given to the 6-mm lens. The results presented in the next section are based on a 6-mm lens.

VARIABILITY IN TRAFFIC EVENT MEASUREMENT

Effect of Data Extraction Method

The method of event time measurement can have a significant effect on measurement precision. For the automated method the video imaging system monitors the video tape recorder's video signal and identifies vehicles by focusing on user-specified detection zones. Each zone consists of a small rectangular grid of light pixels on the video screen. Changes in pixel intensity are analyzed over one or more video frames to determine whether the change is due to the passage of a vehicle. If the change is attributed to a vehicle, the time is then recorded. As a result of this variable-frame processing there is always some uncertainty in the recorded event time relative to the true event time.

The manual method of data extraction is based on a technician's playing back the videotape frame by frame. This method requires considerably more time for data extraction than the automated method because the tape is played frame by frame rather than at normal speed. On the other hand, event times obtained from the manual method can be measured to the nearest frame. As a result, this method has less uncertainty than the automated method.

Headway Measurement

The procedure for measuring vehicle headways by using video images is identical to that which uses observers in the field. A reference line is established on the pavement and on the vehicle. Then the successive vehicular crossing times are recorded and differenced to yield the resultant time headway. The capture of these crossing times with videotape offers several advantages of convenience but, as described below, also reduces the precision of the crossing time measurement.

One source of variability inherent in video-based measurement methods stems from video's discretization of time. The video image is composed of a series of still frames, each of which is analogous to a photograph of events occurring at one instant in time. As these frames are taken every 1/30 sec, manual measurement of event time is limited to an error range of one frame (i.e., ±0.017 sec). Automated measurement of event time has been found to have an error range of two to four frames, depending on the quality of video recording equipment and pavement coloration. The errors from this type of process effectively follow a uniform distribution with a standard deviation of

$$\sigma_F = \frac{n}{30\sqrt{12}} = 0.0096n \tag{5}$$

where σ_F is the standard deviation of measurement associated with frame error (sec), and n is the number of frames in error range (manual, $n = 1$; automated, $n = 2$ to 4).

A number of studies have been conducted by the authors to determine the conditions that dictate the number of frames needed by their automated system (i.e., the Autoscope). These studies indicate that the quality of the video equipment has a significant effect on this number. In particular, professional-grade recording equipment was found to have a range of two to three frames, whereas consumer-grade equipment was found to have a range of three to four frames. Within each of these ranges there is evidence to suggest that the lower number is associated with a light concrete pavement background and the upper value is associated with a black asphaltic pavement background. For subsequent figures presented in this paper, $n = 2.3$ frames was used for the automated method.

A second source of variability stems from the width of the pavement reference line. A wider reference line can increase the uncertainty (i.e., variability) in the resultant estimate of crossing time. The error range is equal to the vehicle travel time across the reference line. The variance of this process can be written as

$$\sigma_W^2 = \text{VAR}\left(\frac{fW_d}{V}\right) \tag{6}$$

where

σ_W^2 = variance of measurement error associated with reference linewidth (sec²),
f = uniformly distributed random variable representing the proportion of the reference line crossed when the event is first identified,
W_d = width of the reference line/detector (m), and
V = vehicle speed [miles per second (mps)].

The methods for computing approximate moments described by Benjamin and Cornell (*3*) can be used to yield the following estimate of σ_W:

$$\sigma_w = \sqrt{\left(\frac{\bar{f}W_d}{\bar{V}^2}\right)^2 \sigma_v^2 + \left(\frac{W_d}{\bar{V}}\right)^2 \sigma_f^2} \tag{7}$$

The bar, or line, over a variable denotes its mean value. For the uniformly distributed variable f with a range of 0.0 to 1.0, $\bar{f}$ equals 0.5 and σ_f equals $1/\sqrt{12}$. In addition, data provided in the *Traffic Engineering Handbook* (*4*) indicate the following relationship between V and σ_v:

$$\sigma_v \approx 0.15\bar{V} \tag{8}$$

Combining these relationships with Equation 7 yields

$$\sigma_w = \sqrt{\left(\frac{0.15W_d}{2\bar{V}}\right)^2 + \left(\frac{W_d}{\sqrt{12}\bar{V}}\right)^2} = 0.30\,\frac{W_d}{\bar{V}} \tag{9}$$

A final source of variability stems from the location of the vehicle reference point relative to the plane of the pavement (which contains the reference line). Specifically, a video imaging system typically locks onto the first part of the vehicle (or its shadow) that crosses the detector. This part is generally the hood, the bumper, or an extended vehicle shadow for the departing, overhead, and approaching views, respectively. Whenever the imaging system locks onto a part of the vehicle that is above the plane of the pavement, the recorded crossing time will be different from the actual time when the vehicle crosses the detector. Heights of 0.9 m (3 ft), 0.5 m (1.5 ft), and 0.0 m were used for the departing, overhead, and approaching views, respectively, in the development of subsequent figures presented in this paper.

Fortunately, the variability associated with an elevated lock-on point can often be avoided. For automated methods this would require an approaching view and sufficient sunlight to create an extended vehicle shadow. For manual methods the camera would be slightly offset from the roadway (as shown in Figure 1) such that the vehicle's tires were in view and the observer would be instructed to use the tire–pavement contact point when the crossing time is recorded.

Based on the geometric relationships shown in Figure 3, the error in crossing time can be computed as

$$e_t = \frac{\delta x}{V} = \frac{L h_{vs}}{V h_c} \tag{10}$$

where

e_t = crossing time measurement error (sec),
δx = distance between the reference line/detector and the vehicle location measured along the center line of the travel path ($= L h_{vs}/h_c$) (m),
L = distance between the reference line/detector and the camera measured along the centerline of the travel path (m);
h_{vs} = height of vehicle signature (i.e., lock-on point) (m), and
h_c = height of camera above plane of roadway (m).

Although this error represents a biased crossing time estimate for any one vehicle, the expected bias for the population of vehicle crossing times is constant and will cancel out when successive pairs of crossing times are differenced to yield the headway. Therefore, the average headway computed from crossing times obtained with an automated method will be an unbiased estimate of the true mean headway.

The measurement variability that stems from locking onto the vehicle above the plane of the roadway can be computed as the variance of e_t. This quantity can be stated as follows:

$$\sigma_{vs}^2 = \mathrm{VAR}(e_t) \tag{11}$$

where σ_{vs}^2 is the variance of measurement error from nonzero vehicle signature height (sec^2).

By the method of approximate moments, σ_{vs}^2 can be estimated as

$$\sigma_{vs} = \left(\frac{L}{h_c}\right)\sqrt{\left(\frac{\bar{h}_{vs}}{\bar{V}^2}\right)^2 \sigma_V^2 + \left(\frac{1}{\bar{V}}\right)^2 \sigma_{hvs}^2} \tag{12}$$

An examination of the relationship between average vehicle hood height and the standard deviation of hood height for 50 vehicles indicated that a trend existed and that the following relationship was reasonable:

$$\sigma_{hvs} \approx 0.15\,\bar{h}_{vs} \tag{13}$$

Combining this relationship and Equation 8 with Equation 12 yields

$$\sigma_{vs} = \left(\frac{L}{h_c}\right)\sqrt{\left(\frac{0.15\,\bar{h}_{vs}}{\bar{V}}\right)^2 + \left(\frac{0.15\,\bar{h}_{vs}}{\bar{V}}\right)^2} = 0.21\,\frac{L\,\bar{h}_{vs}}{h_c \bar{V}} \tag{14}$$

The variance of the event time measurement process is represented by the summation of the previously described component variances. The basis for this formulation stems from the additive nature of component variances of error sources. Similarly, the variance of a headway measurement (i.e., the difference of two event times) is twice that of an individual event time, again because of the additive nature of the variances of a sum (or difference) of random variables. As a result, the combined variance of the three components is

$$\sigma_s^2 = 2\,(\sigma_F^2 + \sigma_W^2 + \sigma_{vs}^2) \tag{15}$$

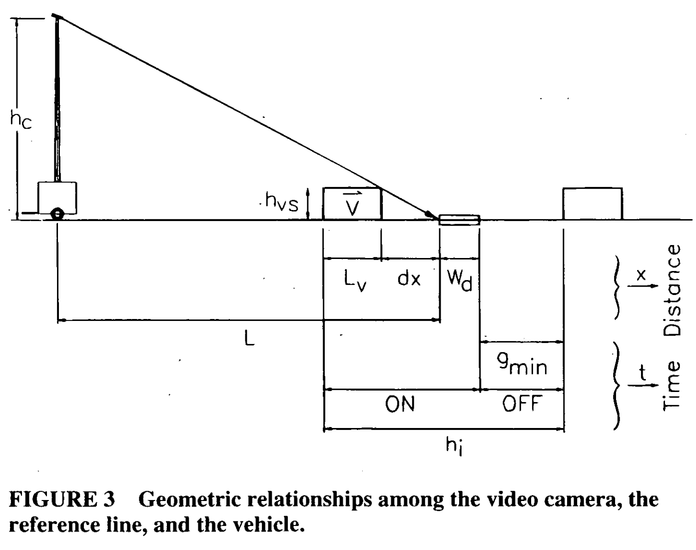


FIGURE 3 Geometric relationships among the video camera, the reference line, and the vehicle.

One final factor that must be considered in estimating headway measurement variance is autocorrelation of the error in successive measurements. An error in measurement of the event time for a subject vehicle affects the headway measured between the subject and preceding vehicle as well as the headway measured between the subject and following vehicle. The resultant measurement error in these two headways is equal but opposite in sign and thus follows a first-order autoregressive process. The implication is that the resultant variability in measurement, as defined by Equation 15, is increased.

Neter et al. (*5*) show that the variance of the error that is due to autocorrelation can be estimated as

$$\sigma^2(\varepsilon_t) = \frac{\sigma^2}{1 - \rho^2} \tag{16}$$

where

$\sigma^2(\varepsilon_t)$ = variance of the error in the first-order autocorrelated random variables,
σ = variance of the error in random error of independent random variables, and
ρ = autocorrelation parameter.

The autocorrelation parameter can be estimated by regressing the measurement error e_h (=true headway − measured headway) for successive headways (e_{hi}, e_{hi-1}). The parameter is equal to the slope of a first-order linear regression model ($e_{hi} = \beta_0 + \rho e_{hi-1}$).

Two tests were conducted to determine the degree of correlation between successive headway measurements. The first test examined several thousand simulated headways, each of which was given a small random measurement error. The headways were assumed to be exponentially distributed, and the measurement errors were assumed to be normally distributed. The autocorrelation parameter from this test was found to be −0.50 regardless of the magnitude of the measurement error variance or average headway. The second test examined several hundred vehicle headways measured with the Autoscope video imaging system. The autocorrelation parameter was found to be approximately −0.45. Giving preference to the latter empirically derived parameter, a value of −0.45 was selected for use in the development of the subsequent figures presented in this paper.

Extending Equation 16 to headway measurement yields the following equation for estimating the variance of the headway measurement process:

$$\sigma_h^2 = \frac{2(\sigma_F^2 + \sigma_W^2 + \sigma_{vs}^2)}{1 - \rho^2} \tag{17}$$

The variables σ_F, σ_W, and σ_{vs} can be obtained from Equations 5, 9, and 14, respectively. This variance represents the error variance that is due to measurement with an imprecise measuring device (i.e., a video-based system). This value would be substituted for s_m^2 in Equation 2.

The standard deviation of the headway measurement process, σ_h, is shown in Figure 4 for a 6-mm lens mounted at a height of 9.8 m (32 ft). As Figure 4 indicates, the standard deviation is lower for higher vehicle speeds because the σ_w and σ_{vs} terms tend toward zero with increasing speed. In addition, the standard deviation increases as the travel direction becomes oriented more toward the camera than perpendicular to it. This increase is due primarily to the effect of view orientation on reference linewidth. The standard deviation is lower when the manual method is used because the event can be measured to the nearest frame (i.e., n = 1 in Equation 5) and because the observer is able to measure the event in the plane of the roadway (i.e., h_{vs} = 0 in Equation 14).

It is worth noting that both the automated and the manual methods yield more precise time estimates than could be obtained by using an external stopwatch during tape replay. Kite et al. (*6*) had a technician observe the video monitor during replay and record (by keyboard toggle) more than 2,300 event times, using a computer-based clock-software package. Each event was measured twice. Kite et al. reported the standard deviation of the difference between these two measurements as 0.4 sec. As a headway is also the difference between two time measurements, a value of approximately 0.4 sec would also be obtained for the standard deviation of headway measurements by this toggle technique. This value is much

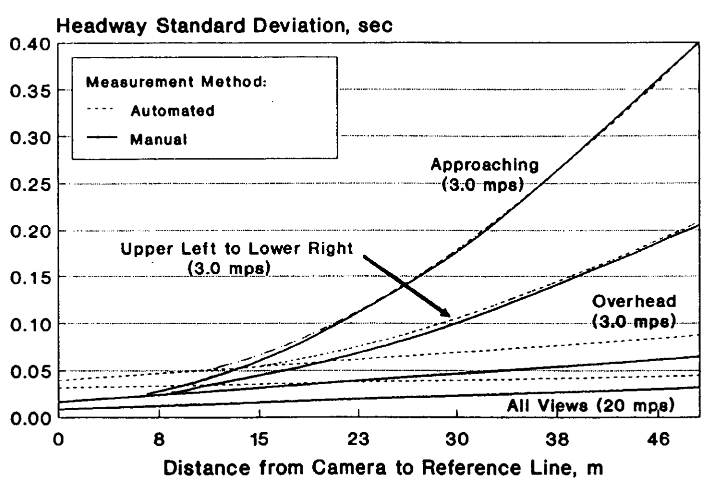


FIGURE 4 Standard deviation of the headway measurement process (1 mps = 3.3 fps).

larger than that obtained from either the automated or the manual method, as shown in Figure 4.

Minimum Measurable Headway

An imaging system typically detects a vehicle by noting when the pixel intensity in the detection zone changes relative to the background or pavement intensity. As a result of this detection technique the system must be able to see the pavement for a minimum duration of time between vehicles to minimize multiple detections (caused by multicolored vehicles or bright reflections) and to refresh its memory as to the background intensity level.

To determine the minimum intervehicle gap time (measured from the back bumper of the lead vehicle to the front bumper of the following vehicle) needed by the imaging system, the gaps between several hundred queued vehicles were measured by both the manual and the automated methods. This study revealed that gaps of less than 0.10 sec were consistently missed by the Autoscope (i.e., it reported two successive vehicles as one long vehicle). On the other hand, gaps of 0.23 sec or longer were always detected by the Autoscope.

The consequence of the imaging system's need for a minimum time gap is that it creates a minimum measurable headway (measured from back bumper to back bumper). This headway can be computed by using the geometric relationships shown in Figure 3 as

$$h_{min} = g_{min} + \frac{1}{V}\left(L_v + \frac{Lh_{vs}}{h_c} + W_d\right) \quad (18)$$

where

h_{min} = minimum headway measurable by imaging system (sec),
g_{min} = minimum intervehicle gap measurable by imaging system (Autoscope: 0.23 sec) (sec), and
L_v = length of vehicle (m).

The minimum headway measurable by the Autoscope 2003 is shown in Figure 5. As this figure indicates, the effect of speed is more significant than view orientation or distance. It also suggests that vehicles moving at low speeds and short headways (e.g., those departing the intersection stop line) may not be individually detectable; rather, they may be detected as one long vehicle. Of course, this operation still allows the Autoscope to function as it is primarily intended (i.e., as a traffic detector for signalized intersection control) because queued vehicle presence is always accurately detected. As speed increases, the minimum measurable headway rapidly decreases to values that are sufficiently short to pose no threat to the precise measurement of individual headways on freeways or expressways.

Travel Time Measurement

The procedure for measuring vehicle travel time by using video images is based on the observation of vehicle crossing times at two reference lines. If the distance between these two lines (in the plane of the roadway) is known, then the speed of the vehicle can be computed. The sources of bias and error variability in this measurement process are similar to those described for headway measurement and are described below.

Travel Time Variability

The variability of measurement stemming from frame error also exists in travel time measurement. As there are two event times measured and differenced to determine travel time, the variability that is due to frame error is additive, as it is with headway. However, direct summation of two quantities of σ_F^2 will overestimate the frame error because of a correlation between the errors of these two measurements (i.e., the errors are not independent). Specifically, the variability in total frame error is reduced somewhat because the same vehicle is being measured at both reference lines. As a result, the measurement errors will be offsetting to some degree. The variance of the error in travel time measurement can be computed as

$$\sigma_{F_t}^2 = \sigma_F^2 + \sigma_F^2 - 2r\sigma_F^2 = 2\sigma_F^2(1 - r) \quad (19)$$

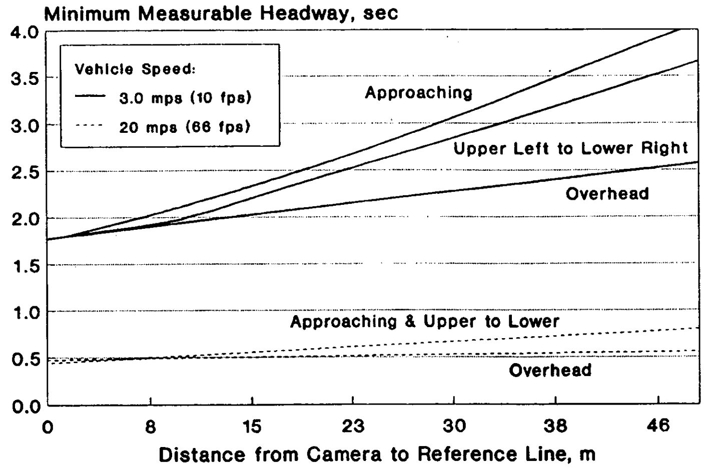


FIGURE 5 Minimum measurable headway with automated data extraction (1 mps = 3.3 fps).

where

σ^2_{Ft} = variance of travel time measurement error associated with frame error (sec²),
σ_F = standard deviation of measurement associated with frame error (Equation 5) (sec), and
r = correlation parameter.

The correlation parameter can be computed by regressing the two event time measurement errors associated with a travel time estimate. The parameter is equal to the slope of a first-order linear regression model. This approach is similar to that described for computing the autocorrelation in headway measurement; however, in this instance the correlation reduces the variability in travel time error.

The correlation between the errors in successive event time measurements for the same vehicle was evaluated by the aforementioned regression approach. This examination indicated that the correlation parameter was approximately +0.35 for both the automated and the manual measurement methods.

The two remaining sources of variability stem from the variability associated with reference linewidth and vehicle signature height. The variance associated with these measurement error sources can be derived from Equations 9 and 14, respectively. As there are two measurements made for each travel time, the total measurement variance for each source can be computed as the sum of the variance of each event time measurement. Thus, the variance of the error that is due to reference linewidth can be computed as

$$\sigma^2_{Wt} = \left(0.30\,\frac{1}{\overline{V}}\right)^2 (W^2_{d2} + W^2_{d1}) \tag{20}$$

where

σ^2_{Wt} = variance of travel time measurement error associated with reference linewidth (sec²),
W_{d1} = width of the first reference line/detector traversed (m), and
W_{d2} = width of the second reference line/detector traversed (m).

In a similar manner, the variance of the error that is due to vehicle signature height can be computed as

$$\sigma^2_{vst} = \left(0.21\,\frac{\overline{h}_{vs}}{h_c\,\overline{V}}\right)^2 (L^2_2 + L^2_1) \tag{21}$$

where

σ^2_{vst} = variance of travel time measurement error associated with nonzero vehicle signature height (sec²),
L_1 = distance between the first reference line/detector traversed and the camera measured along the center line of the travel path (m), and
L_2 = distance between the second reference line/detector traversed and the camera measured along the center line of the travel path (m).

The total variance of travel time error that is due to measurement with videotape systems can be computed as

$$\sigma^2_{tt} = \sigma^2_{Ft} + \sigma^2_{Wt} + \sigma^2_{vst} \tag{22}$$

Travel Time Bias

Vehicle travel time is computed as the difference between two reference line crossing times. However, the automated method can introduce a bias into this travel time computation if the vehicle signature (i.e., lock-on point) lies above the plane of the roadway. With reference to Figure 3, it can be shown that the expected travel time between the reference lines is

$$E[tt] = E\left[\left(t'_2 + \frac{\delta x_2}{V} + \frac{fW_{d2}}{V}\right) - \left(t'_1 + \frac{\delta x_1}{V} + \frac{fW_{d1}}{V}\right)\right] \tag{23}$$

$$= E[t'_2 - t'_1] + \frac{D\overline{h}_{vs}}{\overline{V}h_c} + \frac{1}{2\overline{V}}|W_{d2} - W_{d1}| \tag{24}$$

where

tt = travel time between reference lines (sec),
t'_1, t'_2 = event time measured by the automated method at reference lines 1 and 2 (sec),
D = distance between the two reference lines measured along the center line of the travel path ($=L_2 - L_1$) (m).

The last two terms in Equation 24 represent the average bias in travel time that is due to measurement with the automated method. Both terms are positive quantities, implying that the measured travel time is always shorter than the true travel time. Examination of the terms of this equation indicates that the bias decreases as speed and camera mounting height increase. Alternatively, the bias increases as the distance between the two reference lines increases.

Speed Measurement

Speed Based on Travel Time

An error in the measurement of this travel time translates into an error in the estimate of speed. The method of approximate moments can be used to derive the following equation for estimating the variability of the speed measurement error:

$$\sigma_V = \frac{D}{\overline{tt}^2}\sigma_{tt} \tag{25}$$

where σ_V is the standard deviation of speed measurement error (mps).

The standard deviation of speed measurement is shown in Figure 6 as a function of reference line separation distance D. In the context of speed measurement this distance is referred to as the speed-trap length. As Figure 6 indicates, the uncertainty in measured speed is larger for shorter trap lengths and for automated methods. Depending on the speed being measured and the method of measurement, the uncertainty appears to level off beyond a specific trap length. For low speeds this length is approximately 8 m (25 ft) for both methods. For high speeds using the manual method the length is approximately 23 m (75 ft), and for the automated method it is approximately 46 m (150 ft). These trends are contrary to those noted with respect to travel time error variance because of the travel time term in the denominator of Equation 25. The square of this term tends to be larger and thus more dominant than the σ_{tt} term.

Autoscope Speed Detector

The Autoscope video imaging system has a special detector that is able to measure vehicle speed directly. The use of this detector

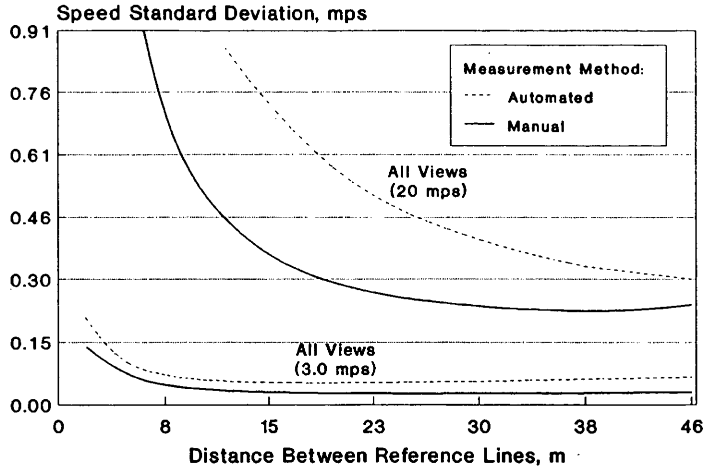


FIGURE 6 Standard deviation of the speed measurement process (1 mps = 3.3 fps).

requires that distances in the plane of the roadway and the height of the camera be input during the calibration step. Once calibrated to the field of view, the Autoscope tracks each vehicle through a short detection zone and estimates its travel speed.

A test was conducted to determine how well the Autoscope's speed detector could measure speed relative to the automated (via travel time) and manual methods. True speeds were measured for 50 vehicles for each of five different fields of view at three different study sites. These true speeds were measured with a tape switch data collection system, as previously described by Moen et al. (7). The camera height was 9.8 m (32 ft). Fields of view included approaching from upper left to lower right, departing from lower left to upper right, and overhead. All views were of the station category. The true speeds ranged from 18 mps (58 fps) to 20 mps (64 fps).

The error, or difference between the true speed and that predicted by the Autoscope, was computed. The standard deviation of this error ranged from 0.84 mps (2.8 fps) to 3.7 mps (12 fps). Comparison of this range with the 20-mps lines in Figure 6 indicates that the range is consistent with the standard deviation obtained with the manual method using 6- to 7-m trap distances. As the length of the Autoscope's speed detector is within this range, this suggests that the Autoscope detector is able to measure speed with the precision of the manual method but with the efficiency of the automated method. However, it also suggests that the precision of the autoscope's speed detector could be improved by increasing its length.

CONCLUSIONS

Time-based traffic events, such as headway and travel time, can be accurately and precisely measured with videotape recording-and-playback systems. The methods of extracting the individual traffic events include (a) manual extraction using a frame-by-frame analysis (using either the in-picture stopwatch image or the frame counter), and (b) automated extraction, using a video imaging system to analyze the tape during playback at normal speed. When data are extracted automatically (or manually by using frame counts), the event times may be biased because of a tape drag effect. The amount of tape drag can vary from system to system, although it is probably lower for systems with professional-grade components.

The automated extraction of vehicle headways may be limited to higher-speed vehicles. The minimum headway that can be reliably detected for vehicles moving at 3.0 mps (10 fps) is ~1.7 sec. The minimum headway for vehicles moving at 20 mps (66 fps) is ~0.5 sec. Manual extraction of headways does not share this limitation.

There are several sources of error variability in the measurement of time-based traffic events by using videotape records. These sources increase the uncertainty in the measurement of an individual event time and, as a result, increase the sample size needed to estimate the mean event time with a desired level of confidence.

One source of error variability in the measurement of event time is due to the discrete time interval between video frames. Another source of error variability stems from the width of the reference line drawn on the video display when it is transformed to the plane of the roadway. A final source of error variability stems from the video imaging system's tendency to lock onto the first part of the vehicle that crosses the detector/reference line. When this part lies above the plane of the roadway, the recorded event time does not equal the actual time when the vehicle reaches the detector's position in the plane of the roadway.

Both automated and manual methods can be used to measure headways, travel times, and speeds. The manual method has lower variability than the automated method in all cases. Thus, the greater precision of the manual method will yield smaller sample sizes than the automated method; however, the automated method can extract data more rapidly. The benefits of each method should be considered on a case-by-case basis.

ACKNOWLEDGMENTS

The authors would like to thank all those who contributed to the development of this paper. They include Brian Moen, Jim Kollbaum, and Mark Lutjeharms, who helped to collect and reduce the field data. A special thank-you is extended to Michael Fowler for his assistance in developing the illustrations used in this paper.

REFERENCES

1. Michalopoulos, P., B. Wolf, and R. Benke. Testing and Field Implementation of the Minnesota Video Detection System (AUTOSCOPE). In *Transportation Research Record 1287*, TRB, National Research Council, Washington, D.C., 1990, pp. 176–184.
2. Doctor, M., and K. Courage. Freeway Management Applications. In *AUTOSCOOP: The AUTOSCOPE User's Newsletter*, University of Florida, Nov. 1993.
3. Benjamin, J., and C. Cornell. *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill, New York, 1970, pp. 180–185.
4. *Traffic Engineering Handbook*, 4th ed. (J. Pline, ed.) Prentice-Hall Inc., Englewood Cliffs, N.J., 1992, Table 2-26, p. 54.
5. Neter, J., W. Wasserman, and M. Kutner. *Applied Linear Regression Models*. Irwin, Inc., Homewood, Ill., 1983, Chap. 13.
6. Kyte, M., W. Kittelson, W. Brilon, and R. Troutbeck. *NCHRP 3-46 Capacity Analysis of Unsignalized Intersections—Phase I Report*. TRB, National Research Council, Washington, D.C., Nov. 1993, p. 39.
7. Moen, B., J. Kollbaum, J. Bonneson, and P. McCoy. Traffic Data Collection Using a Computerized Data Acquisition System. In *Transportation Research Record 1412*, TRB, National Research Council, Washington, D.C., 1993, pp. 39–45.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Development of a Prototype Traffic Safety Geographic Information System

KARL KIM, NED LEVINE, AND LAWRENCE NITZ

A prototype geographic information system (GIS) for the analysis of motor vehicle collisions in Honolulu, Hawaii, is described. An overview of GIS hardware and software is provided along with criteria utilized in the development of the Hawaii system. Mapping and spatial data sources relevant to traffic safety are described and evaluated. The usefulness of Census Bureau Topologically Integrated Geographic Encoding and Referencing (TIGER) files for spatial traffic safety analyses is highlighted. Spatial analyses and potential applications of this technology in traffic safety are also outlined. Recommendations for enhancing the uses of GIS in traffic safety are offered.

In recent years there have been numerous developments in geographic information system (GIS) technologies. Not only is the technology well established but it is here to stay (*1–3*). The emergence of desktop GIS packages has helped to make the technology more widespread, affordable, and accessible. Yet, surprisingly, there have been relatively few accounts of traffic safety programs that use this technology. One exception is a demonstration project developed in Haifa, Israel (*4*). Although GISs are used by no fewer than 29 state departments of transportation, many use these systems solely for engineering drafting and design, and only a few use them for mapping and geographic analysis (*5*). Several developments suggest that this will change. The passage of the Intermodal Surface Transportation Efficiency Act (ISTEA) in 1991 mandates the development of various information systems that could benefit directly from GIS technologies. The federal government has been studying the applicability of GIS technologies in transportation planning for a number of years (*6*). Many states have invested heavily in GIS hardware and software. The GIS has been routinely utilized in transportation planning. More universities are providing training for a cadre of GIS technicians and spatial analysts. GIS technologies have become quite widespread in numerous federal agencies and private-sector organizations (*5*).

In this paper the development of a GIS and spatial analysis system for motor vehicle crashes is described. The system was developed as a tool for research problem identification, and policy formation. In 1992 Hawaii was one of seven states selected by NHTSA to develop a crash outcome data evaluation system (CODES) project. In addition to building data bases containing information on motor vehicle collisions and various health and economic outcomes, the CODES sites were to conduct various analyses on the effectiveness of seatbelts, motorcycle helmets, and other traffic safety interventions. Hawaii's effort included the development of a prototype GIS. Different types of hardware and software as well as sources of spatial data were considered. Geocoding procedures specific to the analysis of collisions were developed.

University of Hawaii at Manoa, Honolulu, Hawaii 96822.

Here various applications of GIS technology in traffic safety are suggested. Recommendations for improving the use of GIS technology are provided in the concluding section.

BACKGROUND OF PROTOTYPE GIS

The prototype traffic safety GIS was developed for Honolulu, Hawaii. Located on the island of Oahu, the city and county of Honolulu consists of a single metropolitan area. With a 1990 population of more than 840,000, Honolulu is one of the largest metropolitan areas in the United States. Honolulu has the largest concentration of population in Hawaii, accounting in 1990 for 75 percent of the state's population and 74 percent of all motor vehicle collisions. In that year there were 19,598 crashes in Honolulu, of which 74 involved fatalities and another 6,733 involved injuries. Problems associated with police crash reports have been widely documented (*7*). In Hawaii, however, there is reason to believe that the quality of police reports is better than in other places. One crash report form is used statewide. In Honolulu all the data are collected by one police department, which uses the standard crash form. Special accident investigators are utilized for collisions involving a fatality or a serious injury. Moreover, the state of Hawaii as well as the city and county of Honolulu have been actively involved in the development and implementation of GIS technology (*8*).

OVERVIEW OF HARDWARE AND SOFTWARE

The GIS is a collection of hardware and software for entering, managing, retrieving, analyzing, and displaying spatial data. Typical hardware includes various input devices (digitizers, scanners), storage devices (disk drives, CD-ROM, tape drives), various processing configurations depending on the operating environment, and a range of different output devices (CRT displays, printers, plotters, and image projection devices). It is often convenient to think of two types of information that can be captured and manipulated with a GIS: graphic and nongraphic, or "attribute," data. The graphic information consists primarily of various elements, such as points, segments or lines, and zones or polygons, that are used in mapping. Attribute data are spatially referenced to the various graphic features and are typically captured and manipulated in a record format by some type of data-base management program and topological algorithm. Together, graphic information and attribute data are especially powerful in traffic safety research. For example, points could be used to identify the locations of events such as head-on collisions or perhaps collisions with blunt-end guardrails. Segments could be used to reference different roadways and the frequency of collisions during peak or off-peak times, or perhaps before and after the installation of lane markers. Zones or polygons can be used to describe or analyze the patterns of crashes that occur within various neighborhoods or travel zones and to relate these patterns to population or employment characteristics.

A GIS should provide the following: (a) information about the location of objects in space (geographical data), (b) information about what occurs at those locations (attribute data), (c) an analysis system for querying features from both the geographical data and the attribute data, and (d) a means for retrieving the information on a display terminal, a printer or plotter, or an external file. Many different GIS programs are available from commercial vendors as well as government sources.

CHOICE OF GIS PROGRAM

GIS programs can be classified as either vector or raster systems. In vector systems, geographical objects are defined by combinations of points, lines (or arcs), and polygons (or areas). Vector programs store the coordinates of the points, lines, and polygons and are programmed to draw lines between the stored coordinates. Attribute data are stored separately but are connected through a data-base management program. Vector systems are efficient in terms of storage and are useful for spatial phenomena that can be meaningfully linked to the basic geometric primitives. For example, motor vehicle crashes are easily linked to a vector system. The crash locations are defined by points, and the streets are defined by lines. Neighborhoods or areas are represented by polygons.

A raster system divides the geographical area into a uniform grid structure. Each cell in the grid is independently coded with regard to an attribute. The geographical and attribute data are intrinsically related. However, every variable that is mapped has to be stored as a separate map layer, although all are referenced to the same grid system. Spatial operations are carried out between the corresponding cells of each layer. There are advantages and disadvantages to raster systems. For spatial phenomena that are continuous and varying (e.g., topography, land uses, soil types), a raster representation is more useful. Operations between layers are more easily conducted for raster than for vector systems; the operations are applied only to individual cells. Raster systems can become very large because the geographical characteristics have to be stored for every data layer. Unlike in a vector system, in which one geographical file can be linked to multiple variables, each data layer represents a separate geographical and attribute file. Storage and processing speed requirements can be demanding with such a system. The major raster GIS programs utilize workstation or mainframe computer technology, for example, Erdas, GRASS, and Vicar. An exception is Idrisi, which runs on a personal computer. Consequently, for a GIS representing motor vehicle crashes a vector system may be more practical for agencies just starting out with this technology.

Another way of distinguishing among GIS programs is to consider whether they are designed to run either on a personal computer or in a workstation or larger computer environment. This distinction between desktop GIS systems and larger GIS programs is more than just a semantic one. Desktop GIS programs are limited by the memory capacity and processing speeds of the personal computers. The complexity and size of the geographic and attribute files are limited. Although large files can be analyzed with a desktop system, there is a price in terms of processing speed. With workstation-based GIS programs, size, complexity, and speed are less important issues, though there are, of course, limits to these systems as well.

There are some other important distinctions between desktop and larger systems. With desktop programs the underlying data structures are much simpler. Most desktop programs, such as ATLAS*GIS, GIS PLUS, and MAPINFO, utilize a simple data structure called a whole polygon structure. With this data syntax, each geographical object (points, segments, polygons) exists and can be manipulated independently of other objects. Changes to the objects (e.g., the addition of a new street) are implemented independently of the existing objects. For a desktop system this is an advantage in that fairly large files can be handled quickly. Most larger systems, however, such as ArcInfo and Intergraph, use a hierarchical data structure in which polygons refer to previously defined lines, which in turn refer to previously defined points. Any change to an object requires a reindexing of all related files in the system, because internal data integrity must be retained. This reindexing requires more data processing steps, more memory, and more processing time. On the other hand, a hierarchical data structure has the advantage of allowing meaningful links to be established among the different objects. For example, building a transportation network that would allow routing decisions to be made requires that all road links be connected and understood by the program; the program has to know that Street A in one part of the map is the same as Street A in another part of the map and also which street links are interconnected. Most of the desktop systems do not have such data structures; they cannot be used as transportation network programs.

The larger, more complex systems are more costly. Most desktop programs cost less than $2,000. Larger systems are frequently leased by the producers at a cost of $5,000 or more per year. Desktop GIS manufacturers can usually provide some data at minimal cost, whereas the manufacturers of larger systems usually do not (most will, however, contract with an agency to construct such files). There is also a user cost. The desktop systems are easy to learn. Typically, there are menus that guide the user through the steps. A traffic safety analyst can become familiar with these systems in a few weeks, sufficiently to be able to develop applications. The larger systems, on the other hand, are typically command driven and are more difficult to use. Because they are typically implemented in shared-user environments, there may be additional networking and system administration requirements. The experience in Hawaii has shown that it takes a user several years to become fully competent with a large system such as ArcInfo. Consequently, there may be a functional difference between a GIS specialist capable of using one of the larger systems and a traffic safety analyst who can use one of the desktop GIS packages.

The criteria used to develop a prototype system for Hawaii involved identifying the various functional capabilities that would be needed. Several key decisions were made early in the effort. First, it was decided that, given the modest resources dedicated to this initiative, existing digital files would be used to the greatest extent possible. The approach was to avoid having to digitize paper maps or to create new base maps from scratch. This also eliminated the necessity for expensive digitizing and scanning equipment. Second, the prototype system had to be working in a matter of months and needed to be functional in terms of both mapping capabilities and analysis. For this reason the simpler desktop packages had a certain appeal. Finally, the GIS work had to be integrated into other project activities. The GIS developed in Hawaii struck a balance between desktop and larger systems. The attribute data were stored and managed in a workstation environment; SAS was the principal software engine for data-base management and statistical analysis. However, the spatial data were manipulated and displayed on a desktop vector GIS using both ATLAS*GIS and MAPINFO.

The setup in Hawaii permits cross-platform and cross-software transfers, thereby taking advantage of the strengths of various packages and sidestepping some of the weaknesses. It is somewhat of an eclectic setup, bringing together conventional GIS packages as well as a number of different specialized tools and utilities. In addition to utilizing SAS for analysis, SpaceStat, a program for zonal analysis (*9*), was acquired. A special software program for analyzing point distributions, called Hawaii PointStat, was written by project staff (*10*).

BUILDING GEOGRAPHICAL FILES FOR REFERENCING CRASHES

A reference system for identifying the location of crashes needed to be constructed. The majority of motor vehicle collisions occur on roads. This necessitated a spatially referenced road file. The Bureau of the Census Topologically Integrated Geographic Encoding and Referencing (TIGER) System was selected (*11*). This is a comprehensive data base of line segments for the entire United States. Each record in the data base represents an individual line segment. An individual record is identified by the type of feature that it represents, for example, a street or a freeway or a railroad line or a pipeline. The TIGER dictionary gives the breakdown of the individual street segments in what is called a census feature classification code.

The data structure of TIGER is topological in that directionality and adjacency are built into the data. Each link has a direction, defined as a line occurring between a "from" node and a "to" node. The nodes are defined by their latitudes and longitudes. There is also a record that defines a shape grammar, which is a separate identifier giving up to 10 additional coordinates for the segment. Because a segment is defined as the line between two nodes, it is graphically drawn as a straight line. However, many segments are not straight but curved. The 1990 TIGER represented an integration of the basic street segment data base (previously called the DIME file) with the U.S. Geological Survey's 1:100,000 quadrangle maps. The scale of TIGER is set at 1:100,000.

A reference system for linking crashes to neighborhoods or small areas was also devised. For several reasons, census geography rather than transportation geography was utilized. First, the census geography is sensitive to residential population. Most trips are homebased, rather than workbased, so a geography relevant to most travel behavior seemed appropriate. Second, there is consistency between one census and another; the census geography is changed only when the population changes considerably. A crash analysis system based on census geography will have continuity for many years. Third, a large amount of data is available for the census geography, particularly from the decennial population and housing census. Fourth, many organizations use census geography and data. Therefore, there is consistency across organizations.

Different levels of census geography, for example, census tracts, block groups, and blocks, were considered. Block groups were selected as the primary spatial units of analysis. A block group is made up of 7 to 15 individual blocks, and a census tract is made up of 3 to 6 block groups. In theory, a census tract should have a population of 3,000 to 5,000, and a block group should have a population of approximately 1,000. In 1990, however, the census geography for Hawaii was too crudely differentiated. On Oahu, for example, for 1990 there were 199 census tracts, each with average of 4,500 persons. However, there were 363 block groups, each averaging slightly fewer than 2,500 persons. Using census tracts would produce too crude a zonal structure. Individual crash locations were aggregated to the block-group level.

Additional geographical files for use in analysis, modeling, or interpretation were generated. All levels of census geography—census tracts, block groups, blocks, and streets—for both the 1980 and the 1990 censuses were constructed. Each of these represents a separate data layer that had to be extracted.

The TIGER files follow the Census Bureau's policy of acknowledging political jurisdictions. In Hawaii this creates special problems because each census unit that borders the coastline actually extends 5 km (3 mi) into the ocean (because this is where the jurisdictional boundary ends). Such maps, although they are correct in terms of boundaries, are not appropriate for analysis of motor vehicle collisions. This problem was corrected by use of a Census Bureau convention that defines the ocean as an individual block and any beach as a separate block. Additional blocks are allocated to harbors, marinas, river outlets, and other features. Deleting the ocean, beach, harbor, and other features produced a reasonably familiar representation of Oahu. Figure 1 shows the block-group structure for Honolulu.

Compared with other mapping sources, such as the U.S. Geological Survey's digital line graph files, the TIGER files are more up to date and comprehensive, particularly for urban Honolulu. Other mapping sources such as U.S. Geological Survey quad maps, the highway inventory, and other spatial files were considered, but problems with referencing census data as well as concerns about scale, accuracy, and accessibility also arose. Because the TIGER files were used to create the base files, data from other sources, such as zone-to-zone trip tables and various trip attractors and generators used in the regional travel demand model, were easily integrated into the system. Data on the location of schools, bars, hotels, tourist sites, and rainfall levels were obtained and used in various analyses. Because all these files are referenced to the same coordinate system (latitude and longitudes in a certain projection), data from additional geographies can be referenced to the baseline block-group geography.

GEOCODING CRASH LOCATIONS

The original motor vehicle accident file did not contain geographic coordinates (latitude and longitude). Crashes were referenced by descriptive information such as the street name and the nearest cross-street name, mile marker, or street address number. Converting the descriptive information into geographic coordinates was done by a semiautomated process known as geocoding.

A standardized dictionary of street names was developed. Crash reports typically list accidents by the intersection nearest to where they occurred (e.g., "at King St. and University Ave."). The sites of freeway crashes are often identified by the nearest freeway exit or entrance (e.g., "H1 at the Vineyard Blvd. off ramp"). However, collisions with only mile marker references could not be automatically matched because milepost locations are not in the TIGER files and are otherwise not available in digital form. Further, other identifiers are often used, such as a street or a bridge crossing (but not actually intersecting) a freeway or the transition road between two freeways, tunnels, or access roads to the freeway. In addition, police reports often use local slang to refer to locations that are not technically identified on a map (e.g., referring to a parallel street as if it were

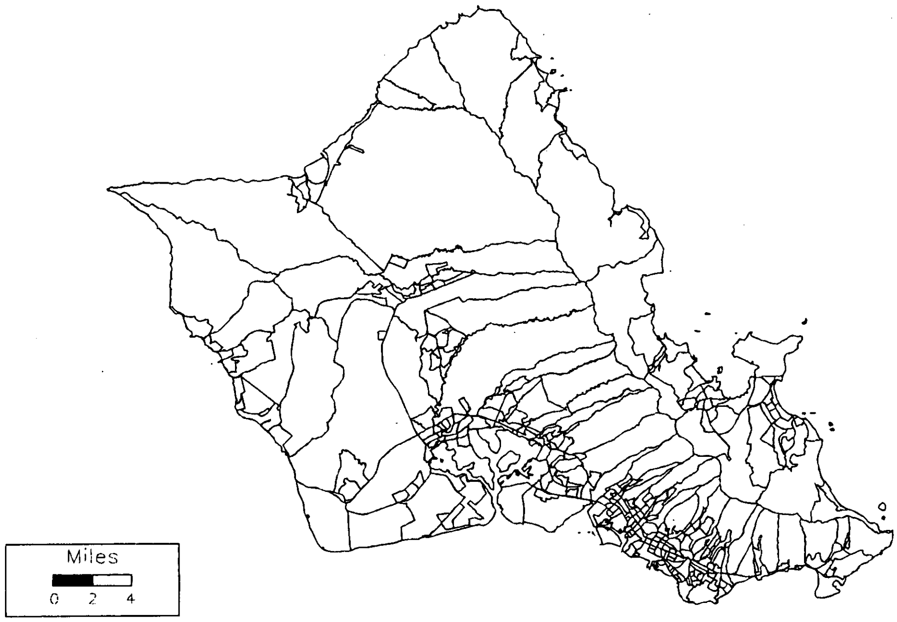


FIGURE 1 Oahu: 1990 block groups (from TIGER files).

intersecting a freeway segment when in fact it does not or referring to a tunnel by the tunnel name rather than by the road name). There are also the usual spelling and typographical errors that affect almost all data bases.

To georeference the crash locations, the 1990 and 1992 TIGER files were used as a basis. The locations were matched to the crash file with a probabilistic linkage software package known as Automatch (*12*). Another program by the same company that developed Automatch, called AutoStan, was used to standardize address references in the crash file for matching against TIGER references. Alternative spellings were included in the dictionary. Artificial intersections had to be created for those crashes for which the reported intersections were not actually intersections (e.g., when a bridge crossed a freeway or when an exit ramp on the freeway accessed a parallel road). Mile marker locations had to be manually processed from paper maps and entered into the dictionary. The process was iterative. It involved, first, a run through all crash locations that were in the dictionary. For those locations that were not identified, specific identifiers were inserted into the dictionary, and the search procedure was rerun. As the dictionary grew, successive codings were able to identify a higher proportion of all names.

Specialized software for matching street names was developed. For every intersection the program would take the standardized TIGER name and search through the street segments to select every link with the name. These would be placed into two matrices, A and B (for each of the two street names). For example, for a collision at King St. and University Ave., all TIGER segments with the name King St. and all TIGER segments with the name University Ave. were selected. Each TIGER segment has two nodes, with a specific longitude and latitude attached to each (a "from" node and a "to" node). The program would compare the longitude of each link in the A list with the longitude of each link in the B list for both nodes (i.e., four separate comparisons for each pair of links from the A list and the B list). For those pairs for which the longitudes in both the A and B lists were the same the program would then compare the latitudes of each pair. Only one node from a single pair would produce a match between the A list and the B list. The longitude and the latitude of the matched node would then be assigned as the crash location. Approximately 2 percent of the crashes had to be identified manually because not all streets in the TIGER system have street identifiers.

The entire process took several months. Approximately 19,213 of the 19,598 locations (or 98.0 percent) on Oahu were identified. For each geocoded crash location, latitude and longitude coordinates were assigned, as were the census tract, block group, and block number for accidents occurring on Oahu. The data were written back into the SAS crash database. Accuracy with this method is reasonable for Oahu. TIGER longitudes and latitudes up to four decimal places were used. The indeterminacy of the fourth decimal place is approximately 6 m (18 ft) [i.e., for latitude, 0.0001 represents approximately 11 m (37 ft); the error of measurement is half of this]. Approximately 43 percent of all collisions occurred at intersections; the error for these is at most a few meters. Of the remaining 57 percent that did not occur at intersections, most occurred on street segments and were assigned to the nearest intersection. For these the error is at most half a block. For freeway collisions, usually the nearest on or off ramp was selected; therefore, 0.02 to 0.4 m (1/8 to 1/4 mi) might be a typical error. Finally, for a few road segments, usually in mountainous or rural areas, assigning the collision to the nearest intersection might produce an error of up to 0.8 m (0.5 mi) or so; there were only a handful of these locations.

ATLAS*GIS and MAPINFO were used for mapping the data. Figure 2 shows the location of the 19,213 crashes on Oahu, each represented by a dot. Since many crashes occur at the same location, the dots are printed over one another. It is clear that crashes are highly concentrated in the built-up urban areas, as the crash pattern overlaps that of urban concentration. Figure 3 shows a detail of the central part of Honolulu, indicating how the crash locations are assigned to the intersections of streets.

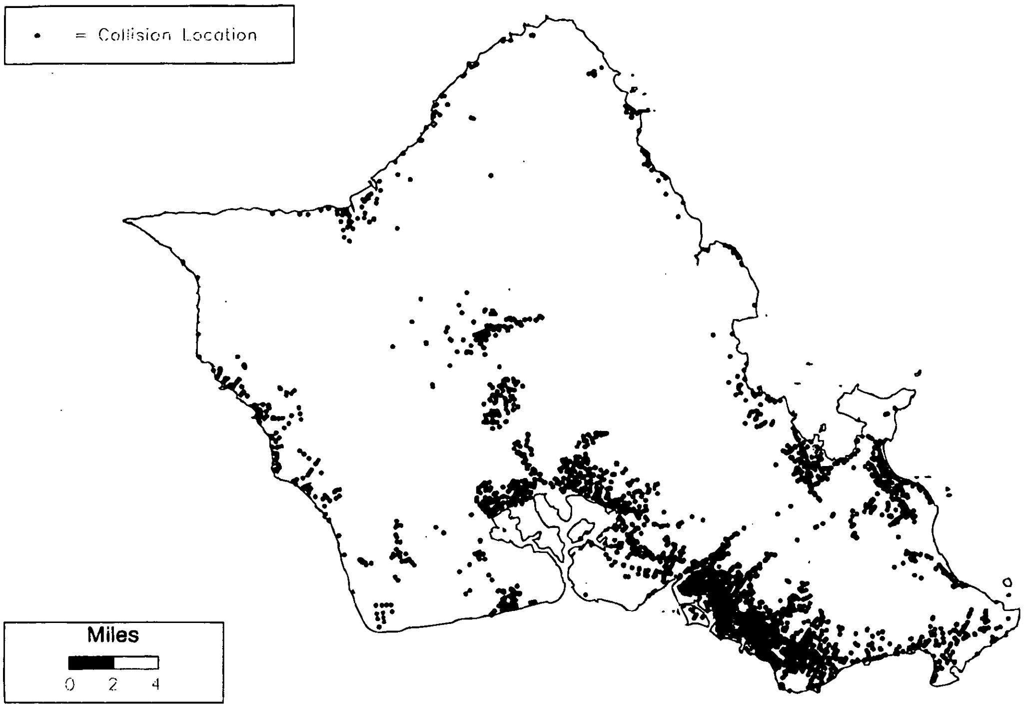


FIGURE 2 Oahu motor vehicle collisions, 1990: location of all collisions.

TRAFFIC SAFETY APPLICATIONS

Many different traffic safety applications were developed. More continue to evolve. The GIS can be used as a tool for description, analysis, and modeling as well as for problem exploration and identification. The information can be used in the formulation of appropriate safety programs.

Based on research facilitated by the GIS, it was determined that the spatial patterns of crashes vary according to time of day and day of week. Presumably these patterns reflect changes in traffic levels. Fatalities and serious injuries were found to have a different spatial pattern from that of noninjury accidents, most likely as a result of differences in speed, alcohol use, time of day (most fatalities occur at nighttime), and home-based activities. Different types of vehicles were found to have different spatial patterns in terms of collisions. The GIS also permitted the examination of relationships between collisions and population size, density, and manufacturing, retail trade, or other employment categories. Moreover, block groups with elementary and junior high schools were found to have a higher share of crashes than would be expected.

Work is also under way to analyze various roadways. Although the study is not yet complete, it appears that freeway ramps and access roads are particularly dangerous locations; block groups with ramps and access roads have much higher likelihoods of crashes when traffic volume and land use are held constant. Freeways by themselves are relatively safe, having much lower accident likelihoods than major arterials, mile for mile. Over time, as more information regarding the roadway inventory is added to the data base, other analyses can be conducted.

The traffic safety GIS has been used to study moped collisions in Honolulu. It helped determine that there are two distinct spatial patterns, one involving out-of-state residents and another involving Hawaii residents. GIS can be a useful tool in identifying special population groups and in locating locations of high incidence of other types of collisions as well. The power of GIS technology is that it lends itself to decision making in terms of spatial features. Certain points, segments, and areas stand out as places for intervention. Intersections (such as those near the University of Hawaii), roadways (Kalanianaole Highway), and districts (Waikiki and downtown) should be examined in more detail to determine the causes and types of traffic conflicts that involve mopeds. A better understanding of the root causes of these collisions could lead to improved policies for enhancing moped safety.

Other applications can be suggested. Areas that have high levels of crashes need to be investigated. There may be unique land uses or activities that generate disorderly traffic flows that, combined with poor signals, traffic routing, or signs, interact to produce dangerous traffic zones.

A traffic safety GIS could be especially useful to community groups in formulating meaningful safety plans. Having a GIS with a visual output can allow groups to develop a common understanding for managing traffic and safety on the neighborhood scale. The availability of maps and spatial data can help to focus discussions among traffic engineers, safety advocates, and concerned citizens.

There are many uses of such a system beyond that which has been described. Such a framework can provide a systematic spatial approach to crash prevention, allowing transportation planners, law enforcement agencies, and other agencies to focus on those areas that have excessively high levels of crashes (beyond that which would be expected). A spatial analysis framework can provide a basis for monitoring future interventions. It can provide a basis for integrating a wide range of information on roadways, vehicles, and drivers.

Another application allows particular crash characteristics to be examined in relation to more general patterns. For example, fatali-

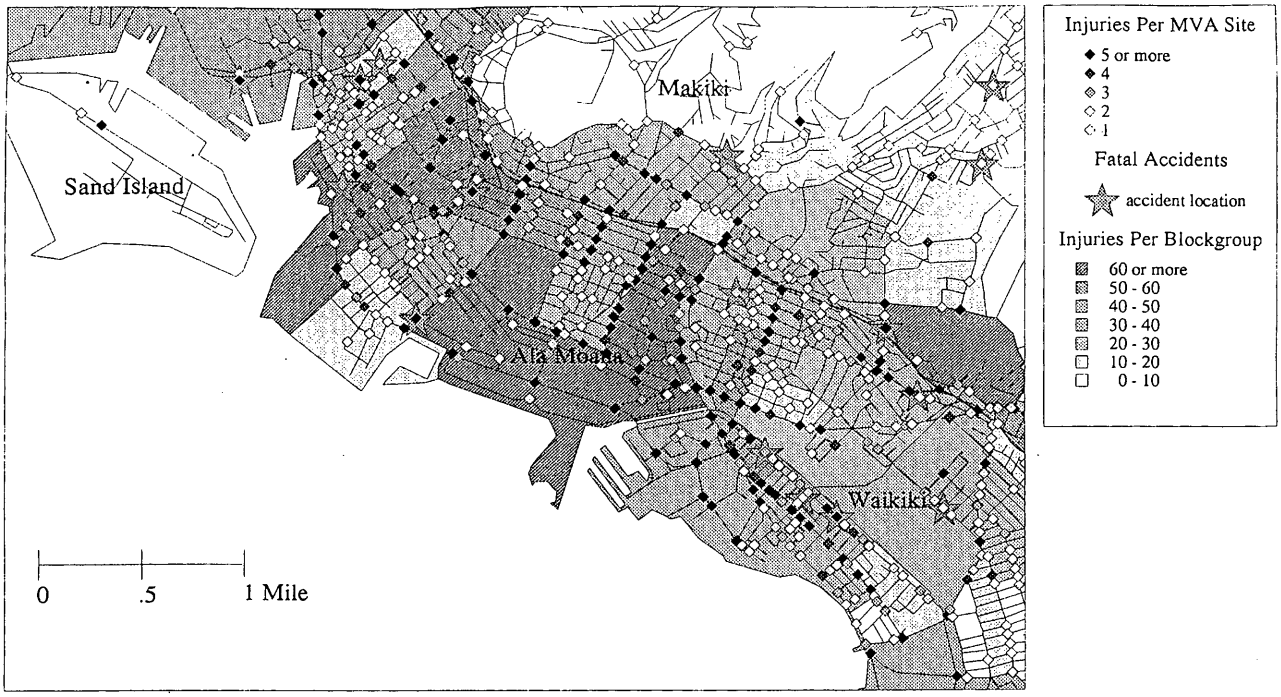


FIGURE 3 Injuries in central Honolulu, 1990 (1 mi = 1.6 km).

ties, alcohol-related crashes, and moped crashes in Hawaii have an essentially different spatial pattern from the majority of crashes. In the case of fatalities there is an interaction among travel speed, alcohol use, and home-based activities. Future research will examine the spatial correlations of crashes with census variables such as unemployment, divorce rates, and income levels. Using the spatial tools developed and utilized here, one can create distinct intervention strategies that target particular populations of vehicle users.

RECOMMENDATIONS FOR FUTURE DEVELOPMENTS

The system developed in Hawaii is only a prototype. GIS technologies have improved greatly in recent years, but there is still need for improvement in terms of making GISs easier to use and developing useful applications of the technology. The problems with a new technology such as the GIS go beyond hardware and software. There are also significant personnel and organizational challenges involving new responsibilities and new approaches to analysis, training, and information management.

The experience in Hawaii suggests that desktop GIS packages could become for traffic safety analysts what spreadsheet programs have become for accountants and others. Although there will always be a need for larger systems, desktop packages provide a powerful tool for research, problem identification, and policy analysis.

At the same time, the availability of desktop GIS packages will not solve all problems. A better system for geocoding crashes is needed, both at the scene of collisions and in terms of processing information from the crash forms. There have been developments integrating Global Positioning Systems and GIS (*6*). Improvements in police crash reporting, especially in terms of geographic referencing, is needed. Perhaps portable field computers can be adapted to contain standard street dictionaries and utilities for improving locational data.

The use of probabilistic matching software such as that utilized in the CODES project can also greatly enhance geocoding and matching of crash locations to existing geographic files (*13*). GIS technologies not only provide a means of satisfying ISTEA mandates regarding the development of various information systems but can also support management decisions regarding new construction, operations, maintenance, and expansion of transportation facilities on information organized, analyzed, and supported with GIS technologies.

There is a need for more geographic standards specific to traffic safety considerations. Although the TIGER files provide a good starting point, there is a need to consider other spatial files and how they can be integrated. Different state and federal agencies need to share and exchange information and procedures relating to the construction of base maps and geocoding of crashes and spatial analyses.

New partnerships need to be forged between federal and state agencies, between universities and departments of transportation, and between private industry and the public sector around the GIS technologies and applications. Above all, there is a need to devote more resources (human and financial) toward meaningful spatial analysis of traffic collisions. By using GIS technologies, perhaps a new generation of analyses can be developed, a situation that can only help to improve the understanding of highway safety.

REFERENCES

1. Maguire, D., et al. *Geographic Information Systems: Principles and Applications.* John Wiley and Sons, New York, 1991.
2. Star, J. *Geographic Information Systems: An Introduction.* Prentice-Hall, Englewood Cliffs, N.J., 1990.

3. Huxhold, W. *An Introduction to Urban Geographic Information Systems*. Oxford University Press, New York, 1991.
4. Peled, A., A. Hakker, and J. Affum. A PC Oriented GIS Application for Road Safety Improvements. *Traffic Engineering and Control*, Vol. 34, 1993, p.3.
5. Attenucci, J., et al. *Geographic Information Systems. A Guide to the Technology*. Van Nostrand Reinhold, New York, 1991, p. 43.
6. *New and Emerging Technologies for Improving Accident Data Collection*. Publication No. FHWA-RD-92-097. FHWA, U.S. Department of Transportation, March 1993.
7. O'Day, J. *NCHRP Synthesis of Highway Practice 192: Accident Data Quality*. TRB, National Research Council, Washington, D.C., 1993.
8. Kim, K. Geographic Systems in Hawaii. In *Geographic Information Systems: Developments and Applications* L. Worrall, ed., Belhaven Press, London, pp. 87–108.
9. Anselin, L., *SpaceStat: A Program for the Statistical Analysis of Spatial Data*. University of California, Santa Barbara, 1993.
10. Levine, N., et al. *Pointstat: A Statistical Program for Analyzing the Spatial Distribution of Point Locations*, Version 1.0. University of Hawaii, Honolulu, 1994.
11. *TIGER/Line Census Files, 1990, 1991*. Bureau of the Census, U.S. Department of Commerce, 1991.
12. Jaro, M. A. *AutoMatch: Generalized Record Linkage System*, Version 2.0. Data Star, Inc, Silver Spring, Md., 1992.
13. Kim, K., and L. Nitz. Application of Automated Records Linkage Software in Traffic Records Analysis. In *Transportation Research Record 1467*, TRB, National Research Council, Washington, D.C., 1994 pp. 50–55.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Changes in the Direction of Urban Travel for the Chicago Area, 1970 to 1990

ED J. CHRISTOPHER, MATTHEW ROGUS, AND SIIM SÖÖT

There is considerable interest in reverse commuting but little understanding of what the term "reverse commuting" means. This study defines directional travel and reverse commuting and examines changes from 1970 to 1990. The household travel surveys of the Chicago Area Transportation Study are used to show that work trips are still very strongly oriented toward the central business district (CBD). Even though there have been increases in the number of workers commuting toward peripheral locations, the proportion commuting toward the CBD has remained fixed at 52 percent or has declined only slightly, depending upon how the CBD is defined. Work and shopping destinations illustrate an expected geographic pattern. Movement from the central county (Cook County) is more likely to be outbound than that which originates in peripherally located counties. Work trips show the strongest directional bias toward the CBD, whereas both recreational and shopping trips have a slightly greater tendency to be directed away from it. The study concludes that all modes of travel, except by suburban bus and school bus, illustrate a bias toward the CBD despite rapid population decentralization and an increasingly diffuse pattern of urban development. Further, these directional biases have changed little in 20 years, providing insights into the interaction between how a city grows and how the travel behavior of its populace changes.

During the past half-century American cities have grown territorially, changing from core-dominated to dispersed low-density regions. Whereas transportation systems previously focused on travel to the downtown, they now must provide access to many diverse locations. For many travelers, however, the destinations away from home are still directed toward the downtown. Studies have shown that there exists a sectoral (radial) bias in daily trip activity and that the majority of trips from home are toward rather than away from the central business district (CBD). But, considering the evolution of American cities, the decentralization process has been so strong that the traditional wisdom regarding trip direction and sectoral bias may no longer hold.

This study will examine changes in urban travel activity and determine to what extent the direction and restructuring of travel has changed with urban decentralization. This study is important because it (a) provides useful information about the travel behavior that contributes to inbound congestion; (b) identifies the degree of outbound travel activity, which is commonly not well served by public transportation; (c) shows the changes in directional travel over a 20-year period; (d) provides a discussion of reverse commuting and reverse travel; and (e) suggests a method of defining reverse travel that is transferable to other areas.

The travel data examined here will provide insight into how travel behavior changes as urban form evolves. It will address two related questions: Is there proportionately more reverse commuting in contemporary metropolitan areas, such as Chicago, and is the increase a product of growing labor force participation rates and population growth?

E. J. Christopher and M. Rogus, Chicago Area Transportation Study, 300 West Adams Street, Chicago, Ill. 60606. S. Sööt, Department of Geography, MC 183, University of Illinois at Chicago, Chicago, Ill. 60607-7138.

BACKGROUND

Reverse commuting has been observed for several decades (*1*), and it continues as suburban employment opportunities grow. To assess the magnitude of this reverse travel activity it is necessary to establish a definition of reverse commuting. In its simplest form reverse commuting can be defined as work trips from the central city to suburban job locations (*2*). It may also have a temporal dimension, that is, a peak directional demand. However, there are two problems with the city-to-suburb definition. First, it is not uniformly accepted. A recent study for the Chicago Area Regional Transportation Authority (*3*) identified the reverse commute as being from city to suburbs, but for commuter rail operations the "city" included three high-density suburbs—Evanston, Cicero, and Oak Park. However, the city is not consistently confined to a few suburbs but may include all close-in suburbs (*4*).

Second, if only central-city residents were eligible to be reverse travelers, then the definition would be totally dependent on the territorial size of the central city. In the United States the central-city size varies from less than 50 mi^2 (130 km^2) for municipalities such as Boston, Buffalo, Miami, and San Francisco to more than 300 mi^2 (780 km^2) for Dallas, Houston, Indianapolis, Kansas City, Los Angeles, New York, Phoenix, and San Diego. Consequently, many of these areas lack an ideal fit between the old established, high-density part of the metropolitan region and the city limits. Therefore, the city-to-suburb definition yields very different travel patterns in Boston and Houston and restricts interregional comparisons.

To remedy this, a definition based on the prevailing direction of movement was established. It is patterned after the definition used by Marston et al. (*5*). This definition is centered around one key question: Is the trip headed in a direction toward or away from the CBD? In as much as travel from any location in the metropolitan area toward its CBD is likely to be targeted at higher-density destinations with a more established street network and a greater density of public transit coverage, it is very different from a trip toward the urban periphery. However, there will always be some trips that are neutral, in which the destination is not clearly closer to or farther from the CBD. This neutral category includes two types of trip: local trips, which are less than 1 mi (0.3 km) and longer trips, called lateral trips (Figure 1). Each trip is placed into one of four categories:

1. **Toward the CBD:** Trips toward the CBD are those for which the destination is at least one airline mile closer to the CBD than the trip origin.

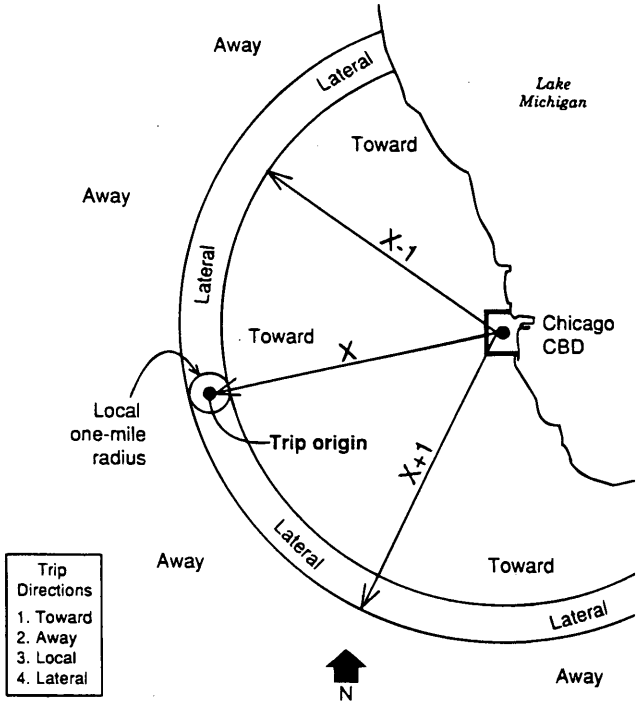


FIGURE 1 Four trip-direction categories.

2. **Away from the CBD:** Trips away from the CBD include those for which the destination is at least 1 mi farther from the CBD than the trip origin. Trips in this category are classified as reverse commutes.
3. **Local (Neighborhood) Trips:** Neighborhood trips have destinations within a 1-mi radius of the origin.
4. **Lateral Trips:** Lateral trips have destinations outside the neighborhood but are not more than 1 mi farther from the CBD or less than 1 mi closer to the CBD than the trip origin. These trip definitions are shown graphically in Figure 1.

To apply these definitions a point within the CBD must be identified that represents the focal point of the downtown. In Chicago, the intersection of State and Madison streets was chosen, which incidentally is the origin of the city's address system.

Before continuing further it is necessary to discuss how the definition may be applied to a typical sequence of trips. Figure 2 illustrates a work trip chain in which there are five destinations and five trips. There are at least two ways to interpret trip directions: unlinked trip directions (Figure 2, *top*) and home-based directions (Figure 2, *bottom*).

In the case of unlinked directions the directionality definition shown in Figure 1 is applied from the origin to the destination of each trip segment of the trip chain. For example, the eat-meal trip before the work trip will have the direction specified by arrow 1 in Figure 2 and is classified as toward the CBD. Trip 2, to work, is short (less than 1 mi) and is classified as local, and the shopping trip (arrow 4) is away from the CBD. In this scenario the directional criterion is reapplied to each additional origin.

For home-based directions the directionality definitions are applied to each destination, holding the origin constant at the home location. The interconnections between destinations (the actual trips) are not considered except for the first trip. Note that in a chain of N stops there are $N - 1$ home-based directions, and, in this example, all four are toward the CBD.

In effect, the unlinked direction considers both the trip and its direction, whereas the home-based direction focuses only on direction. For the transportation planner this definition of home-based direction should not be confused with the traditional definition of home-based trips.

PURPOSE OF STUDY

The focus of this paper is on both the unlinked and the home-based directions shown in Figure 2. The unlinked direction is useful because it traces the actual travel route. On the other hand, the home-based analysis yields information about where the destinations are in relation to the home location, thus providing insights into travel in the context of urban form. This paper, then, contributes background information to the debate on jobs-housing imbalance (*6,7*) by illustrating that little has changed in directional travel in 20 years.

The home-based direction indicates to what extent urban residents continue to relocate toward lower-density neighborhoods from their places of work. Larger homes or real estate parcels can frequently be purchased in neighborhoods farther from the CBD or lower housing costs are achieved through the trade-off with transportation costs (*8*). Although the housing and transportation costs trade-off has been theorized for decades, evidence shows that urban growth can nullify this relationship (*9*). With both residences and jobs decentralizing, the locational advantage of living near the downtown begins to dissipate. Further, many blue-collar workers living in the urban core cannot take advantage of white-collar jobs in the CBD and have moderately long reverse commutes.

This paper contributes to the discussion of both the jobs-housing imbalance and the housing-transportation cost trade-off by examin-

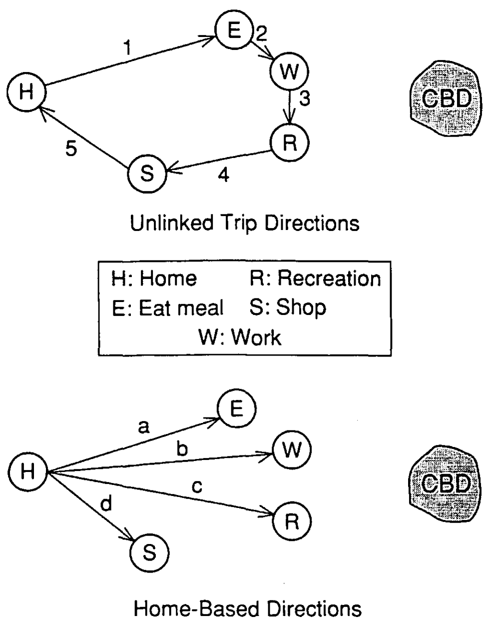


FIGURE 2 Unlinked and home-based trip directions.

ing a very specific aspect of travel behavior, travel direction and its change over time. This includes an examination of what proportion of travel is also directionless, whether by being too short (local trips) or by having no discernible direction (lateral trips). These two travel categories are examples of trips that do not typically support either the imbalance or the trade-off argument.

DATA AND STUDY AREA

The study uses the household travel surveys conducted by the Chicago Area Transportation Study (CATS) in both 1970 and 1990 (*10,11*). The 1970 home-interview survey includes information on 21,748 households and 113,653 trips. The 1990 study includes 19,314 households and data on 162,755 trips.

There are differences in the data sets. For example, the 1970 data collapsed access trips for commuter rail into the traditional priority mode trips. But the 1990 data contained access and egress modes as separate trips. For this analysis the 1990 data base was adjusted so direct comparisons could be made.

There are other more subtle differences in the two data sets. The most notable is the change in method from a home-interview format to a mail-out mail-back survey. Another change for the 1990 data is the inclusion of more trip purposes, such as eating out and banking. This change led to the elimination of the personal business category, which has traditionally plagued surveyors as being difficult to define.

The choice of a 1970-to-1990 comparison is advantageous for several reasons. First, the late 1970s and early 1980s can be shown to be highly unstable in terms of urban development, maturation, and growth. Second, 1980 would reflect the high energy costs in the early 1980s and the energy shocks experienced by the traveling public in the late 1970s. In 1981 the average cost of a gallon of gasoline was approximately $1.90 (using 1994 dollars), in contrast to less than $1.15 per gallon in early 1994. Third, the population driver, the Baby Boomer, moved from a prefamily to a mature-family stage, facing expenses such as college tuition during this period (*12*).

The study area consists of what is generally known as Northeastern Illinois or the metropolitan Chicago region. It includes six counties, with the addition of Kendall County in the 1990 data. Because Kendall County has a population of only 38,000, it represents well under 1 percent of the population of the entire study area. Therefore its addition to the data does not significantly affect the comparison of the 1970 and 1990 data. It should be noted that part of the analysis focuses on these geographic subdivisions or counties. Cook County is further divided into three subareas, including suburban Cook County, the city of Chicago, and the geographic CBD (CBD zone).

Between 1970 and 1990 the region's population grew by approximately 4 percent to just over 7.1 million, but the urban territory expanded by nearly 50 percent. One measure of this population decentralization is the size of the region's Census Transportation Planning Package. The population of the 0.25-mi^2 (0.65-km^2) traffic analysis zone has increased from 4,820 in the 1970 package to 10,060 in the 1990 package. Unlike many other metropolitan areas, the Chicago area has substantially expanded geographically, with little population growth.

FINDINGS

Patterns in 1990

The regional variations in trip direction are examined in categories of purpose and mode. Each of these trip types has a distinct set of directional patterns.

Trip Purpose

For any home location, work and work-related destinations have the highest propensity to be headed toward the CBD, in both cases just over 50 percent (Table 1). Conversely, just under one-fourth of these trips are to locations farther from the CBD than the home and would be considered reverse commutes. In effect, the reverse component of all work trips is clearly in the minority. Beyond that, it is not perceptibly different from the reverse direction for other purposes (Table 1). With the exception of passenger and banking trips, which are largely local activities, the proportion of away trips ranges from 22 to 26 percent. For purposes such as banking, eating a meal, and serving a passenger, the trip toward the CBD is the prevailing one. None of the purposes is found in greater proportions heading away from the CBD. Both shopping and recreation trips have the best balance between trip ends toward and away from the CBD.

Except for work and work-related activities, most other travel purposes are dominated by short trips. Banking, serving a passenger, and shopping all have more than a third of their destinations less than 1 mi from home. Perhaps only banking shows prospects for vehicle trip reduction, provided that it is not chained to other purposes. Shopping and serving a passenger are likely to require a vehicle.

The most consistent direction is lateral. This category includes trips with destinations outside the local neighborhood that show no substantial progress toward or away from the CBD. From 11 to 16 percent of trips, regardless of purpose, fall into this lateral category.

TABLE 1 Trip Direction by Purpose for 1990 (Home-Based Direction)

Trip Purpose	Trips (mil)	Toward CBD	Away from CBD	Local	Lateral
Work	3.56	52%*	24%	14%	11%
Work related	0.76	51%*	24%	13%	12%
Serve Pass	1.40	27%	17%	41%*	16%
Other	1.67	30%	23%	33%*	14%
School	0.71	30%*	26%	29%	15%
Recreation	1.05	29%	26%	30%*	14%
Return Home	0	NA	NA	NA	NA
Bank	0.35	28%	15%	42%*	15%
Eat Meal	0.79	39%*	23%	25%	13%
Shop	2.05	24%	22%	38%*	16%
All Trips	12.34#	39%*	21%	27%	13%

* Highest values for trip purpose.
# Total of 14.34 million directions; 2 million change-of-mode trips not included.
Source: Computer from the 1990 CATS Household Travel Survey.

Work Trips Travel to work creates the greatest stress on the highway system, in part because the travel is very directionally biased. Figure 3 illustrates the regional variations in unlinked trip directions. It again underscores the difference between inbound and outbound commuting found in Table 1 (in which the home-based direction was examined). Four of the most peripheral counties have the highest amount of directional travel heading toward the CBD. For each one, more than half of all the work trips are destined toward the CBD. Even in Chicago, without considering those who live within the geographic boundary of the CBD, 43 percent of travelers commute toward it, only slightly fewer than those in suburban Cook County (46 percent) and DuPage County (48 percent).

Conversely, more than a fourth of the residents of the city of Chicago commute in an outbound direction. This figure drops from 27 percent in the city to 23 percent in both suburban Cook and DuPage counties. In effect, the outbound commuting behavior is not unique to the city but is almost equally common in the inner ring of suburbs. As little employment is found in the peripheral areas in the distant counties other than Cook County, the outbound percentages drop into the teens.

Not surprisingly, short-distance commuting (less than 1 mi) declines from 44 percent in the CBD zone and 21 percent in the city of Chicago to less than 20 percent in the first ring of suburbs. It is back up in the 20 percent range in the three western most counties. One explanation for this is that in the higher-density parts of Chicago job opportunities are more prevalent, and again in the most distant suburbs the labor force is believed to be rather localized.

One of the most uniform patterns is the consistency in the local and lateral categories, which together identify a lack of directional trends. For most of the study zones 30 percent of trips show no distinct directionality. Two clear exceptions are the CBD zone and Kane County, both of which have many local job sites. In the case of Kane County, its principal north-south urban corridor is along the Fox River Valley, which is largely equidistant from the CBD.

Considering that the private vehicle is so dominant in the work trip, the pattern for drivers was also examined and found to be rather similar to Figure 3. Because of that similarity it has not been included as a separate figure. However, the major difference is a slightly higher degree of outbound commuting. In the city of Chicago the away direction accounts for 39 percent of the private vehicle work trips, and the toward-CBD direction constitutes 37 percent. These percentages indicate that the outbound movement by drivers is greater but only by a small margin.

Shopping Trips Unlinked shopping trips have a very different directional pattern. Unlike work trips, shopping travel is largely local. A shopping trip distance of less than 1 mi ranges from 30 percent in the low-density counties of Will and Kendall to 49 percent in the city of Chicago and 59 percent in the CBD zone (Figure 4). As a whole, 40 percent of all shopping trips are less than 1 mi long, in contrast to only 20 percent for work trips.

The strongest shopping directional bias occurs in the city of Chicago, where 24 percent of the trips are outbound and only 15 percent are inbound. Suburban Cook County also has a stronger outbound than inbound movement, as does DuPage, although the latter is not statistically significant. Looking back to Figure 3, one finds that there are no such examples for work trips. In sharp

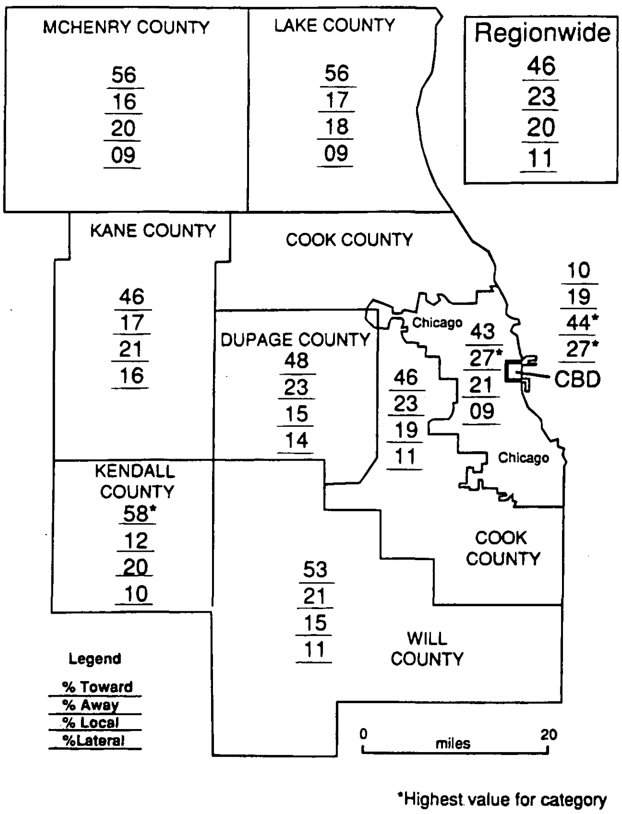


FIGURE 3 Direction of unlinked work trips for 1990.

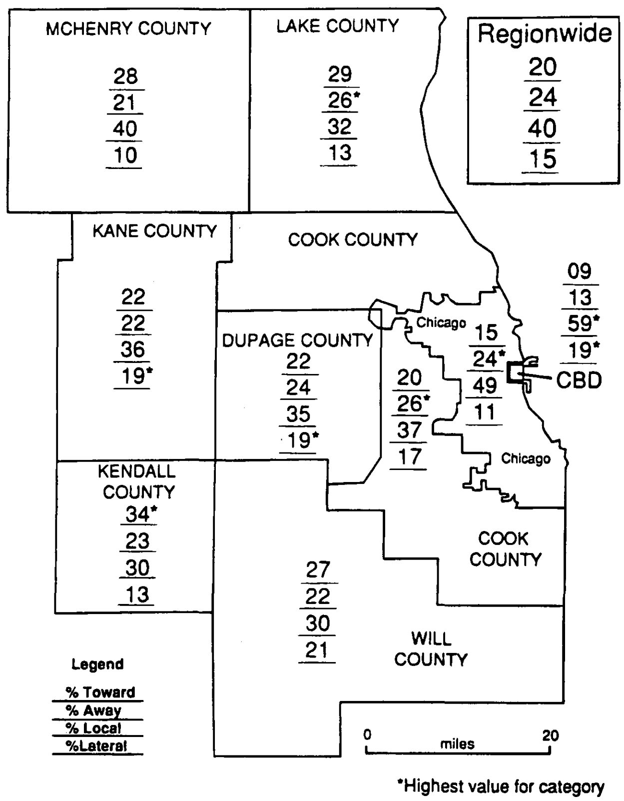


FIGURE 4 Direction of unlinked shopping trips for 1990.

contrast, inbound work trips outscore outbound trips by at least 15 percentage points for all study zones.

The outbound movement from the city of Chicago reflects travel to suburban shopping centers, where parking is easy and the diversity of merchandise is great. Particularly popular are suburban shopping centers on radial expressways leading from the city. A survey of municipal stickers by one of the authors found that as many as one-third of the vehicles at a popular suburban shopping center had Chicago vehicle registrations.

In sum, the outer counties tend to shop toward the CBD, and the inner areas shop in directions away from the CBD. All places have a large local shopping component. As such, shopping trips do not compound the CBD directional biases found in work trips.

Travel Mode

An analysis of travel direction by mode shows that the results are very different for the home-based direction than for unlinked directions. If one examines the 24-hr day by using the unlinked directional criteria, then, as Table 2 shows, the directions to and from the CBD are largely in balance. Table 2 is not therefore particularly revealing except that it illustrates the radial nature of the two rail modes, which exhibit only a few trips to local and lateral destinations. The local category also reflects the average trip length of these modes. For example, 43 percent of the taxi trips are less than 1 mi. More informative was the examination of directionality measured during the 6:00 a.m. to 10:00 a.m. peak period. Here it was found that within the individual regions represented in Figure 3, 92 to 100 percent of the commuter rail trips were headed toward the CBD. Regionwide 95 percent of the commuter rail trips headed toward the CBD, whereas 4 percent headed away.

Considering the location of the destination using the home-based direction definition (Table 3), it becomes apparent that, besides the taxi mode, the two rail modes, commuter rail and subway–heavy rail, have the strongest CBD orientation. However, there is a substantial difference in destination for riders of city and suburban buses. In Chicago there is a definite CBD directional bias, but in the suburbs there are more trips away from the CBD than toward it—39 percent versus 32 percent. Clearly, the suburban bus system provides an important function, serving a large number of riders in the reverse direction.

Both automobile drivers and passengers are more likely to be destined toward the CBD than away from it. Passengers frequently make short trips, and because 26 percent have trips of less than 1 mi (local trips), the proportion destined toward the CBD is only 32 percent but it is still higher than the 27 percent for trips that are away from the CBD.

Perhaps the most important finding between the two directional definitions is in the walking mode. Whereas more than 90 percent of the walking trips are less than 1 mi long (Table 2), only 41 percent of the walking trips are made close to home (Table 3). Of those trips whose destinations are not close to home, the overwhelming proportion are made closer to the CBD than away from it in reference to the home location. In effect, then, if a person's work trip is toward the CBD, he or she is much more likely to walk to work or to be out walking than if the work location is farther from the CBD than from home.

TABLE 2 Trip Direction by Mode for 1990 (Unlinked Trip Direction)

Trip Mode	Trips (mil)	Toward CBD	Away from CBD	Local	Lateral
Walk	1.59	2%	2%	91%*	4%
Auto Drive	13.11	29%*	29%*	28%	15%
Passenger	2.26	27%	27%	28%*	17%
Suburban Bus	0.09	33%*	33%*	18%	16%
Commuter RR	0.33	49%*	49%*	1%	0%
City Bus	1.16	34%*	34%*	17%	16%
Subway/El	0.55	46%*	46%*	3%	4%
School Bus	0.18	27%	30%*	20%	23%
Taxi	0.08	23%	22%	43%*	15%
Other	0.16	22%	22%	43%*	15%
All Trips	19.51	27%	27%	31%*	14%

* Highest values for trip mode.
Source: Computed from the 1990 CATS Household Travel Survey.

TABLE 3 Trip Direction by Mode for 1990 (Home-Based Direction)

Trip Mode	Trips (mil)	Toward CBD	Away from CBD	Local	Lateral
Walk	2.05	50%*	5%	41%	4%
Auto Drive	8.57	36%*	24%	26%	14%
Passenger	1.48	32%*	27%	26%	16%
Suburban Bus	0.10	32%	39%*	15%	14%
Commuter RR	0.34	60%*	6%	23%	11%
City Bus	0.91	46%*	21%	18%	15%
Subway/El	0.58	54%*	10%	29%	8%
School Bus	0.11	20%	36%*	22%	23%
Taxi	0.07	66%*	11%	14%	9%
Other	0.13	32%*	25%	29%	14%
All Trips	14.34	39%	21%	27%	13%

* Highest values for trip mode.
Source: Computed from the 1990 CATS Household Travel Survey.

Comparison of Unlinked and Home-Based Directions

The data for the two directional definitions are generally similar, but there are some notable differences. The local category for the work trips illustrates one difference. Table 4 shows that although 14 percent of the work trips are to job sites within 1 mi of the home, 20 percent of the airline trip lengths to work (measured from the last stop) are 1 mi or less. This difference is even greater for work-related activities. Table 4 suggests the obvious, that these work-related trips have destinations that are closer to work than to home.

Shopping is also illustrative of the differences. Whereas 28 percent of the shopping activity takes place within 1 mi of the home, 40 percent of the shopping trips are less than that length. Concerning the CBD zone, shopping trip directions tend to be away from the CBD, perhaps on the way home from a city work location. However, when shopping trip direction is evaluated in reference to home locations, more shopping destinations are closer to than farther from the CBD.

In general, the home-based direction suggests a stronger "toward the CBD" bias than the unlinked direction and is more directly linked to urban form.

Changes from 1970 to 1990

During the 20-year period there were dramatic changes in land use and travel behavior. Vehicular ownership increased dramatically (*13*), and the region grew territorially (*14*). Against this backdrop one would expect the reverse traffic to decrease with population decentralization but to increase with employment decentralization and, because of the greater dispersion, for local activity to decline. How these forces balance is the subject of this section. This examination also adds to the debate regarding the travel efficiencies of core-dominated versus dispersed metropolitan regions (*15,16*). The Chicago area was clearly much more dispersed in 1990 than it was in 1970.

Despite major decentralization forces there was little change in the directionality of trip purposes. There were moderate declines in trips toward the CBD for work, shopping, and recreation, but recreation also exhibited a decline in trips away from the CBD (Table 5). Also, although work trips shifted away from the CBD, work-related travel showed a slight shift toward it.

Work Trips

There was little change in directional activity for the trip to work (Table 5). Many of the four directional categories registered no more than a two-percentage-point shift. Although the unlinked directional analysis for work trips exhibits a slight shift from the direction toward the CBD to the direction away from it, the ratio of toward versus away is still 2:1. In contrast, the home-based direction stayed at 52 percent toward the CBD for both 1970 and 1990. The away direction increased from 21 to 24 percent, and the other two directional categories declined slightly.

TABLE 4 Trip Direction by Purpose for 1990 [Unlinked Trip Directions (UT) and Home-Based Directions (HB)]

Trip Purpose	Toward CBD		Away from CBD		Local		Lateral	
	HB	UT	HB	UT	HB	UT	HB	UT
Work	52%*	46%*	24%	23%	14%	20%*	11%	11%
Work Related	51%*	34%*	24%	27%	13%	26%	12%	12%
Serve Pass	27%	25%	17%	21%	41%*	39%*	16%	16%
Other	30%	25%	23%	25%	33%*	36%*	14%	14%
School	30%*	25%*	26%	26%	29%	32%*	15%	16%
Recreation	29%	25%	26%	28%	30%*	33%*	14%	15%
Return Home	NA	22%	NA	34%*	NA	30%	NA	14%
Bank	28%	22%	15%	19%	42%*	46%*	15%	14%
Eat Meal	39%*	20%	23%	21%	25%	44%*	13%	15%
Shop	24%	20%	22%	24%	28%*	40%*	16%	15%
All Trips	39%*	27%	21%	27%	27%	32%	13%	14%

* Highest values for trip purpose.
NA Not applicable.
Source: Computed from the 1990 CATS Household Travel Survey.

TABLE 5 Changes in Trip Directions by Purpose from 1970 to 1990 (Unlinked Trip Direction)

Trip Purpose	Toward CBD		Away from CBD		Local		Lateral	
	1970	1990	1970	1990	1970	1990	1970	1990
Work	49%*	46%*	22%	23%	17%	20%	12%	11%
Work Related	31%	34%*	32%*	27%	24%	26%	13%	12%
Serve Pass	24%	25%	16%	21%	41%*	39%*	19%	16%
Shop	23%	20%	22%	24%	38%*	40%*	18%	15%
Recreation	28%	25%	29%*	28%	27%	33%*	17%	15%
Bank	NA	22%	NA	19%	NA	46%*	NA	14%
Eat Meal	NA	20%	NA	21%	NA	44%*	NA	15%
Personal Bus	31%	NA	23%	NA	30%*	NA	16%	NA
Other	NA	25%	NA	25%	NA	36%*	NA	14%
Return Home	23%*	22%	35%*	34%*	26%	30%*	16%	14%
All Trips	28%	27%	29%*	27%	27%	31%*	16%	14%

* Highest values for trip purpose.
NA Data not collected.
Source: Computed from the 1970 and 1990 CATS Household Travel Surveys.

Table 6 shows that driving toward the CBD has declined, but the two-percentage-point drop does not identify a statistically significant change. Besides trips in the Other category, the largest drop in commuting toward the CBD is for automobile passengers or ridesharers, from 42 percent to 37 percent. Still, there is no corresponding increase for commuting away from the CBD; the increase is for local ridesharing.

An examination by county of trip origin shows that for the central county (Cook) there was a substantial change in the commute toward the CBD and in local commuting. However, as these offset each other, there was little change in commuting away from the CBD (Figure 5). Work trips away from the CBD increased from 23 percent in 1970 to 25 percent in 1990, only a two-percentage-point change. In contrast, local work trips increased from 16 to 21 percent, a five-percentage-point increase.

Commuting toward the Chicago CBD also declined for DuPage County but showed a strong increase for most of the peripheral counties. McHenry, Kane, and Will counties all showed approximately a 10-percentage-point increase in the commute toward the CBD. In many of these cases, local work trips declined as a proportion of all work trips, and the gains were toward the CBD.

Shopping Trips

Table 7 shows only minor changes and shifts for automobile shopping trips. Local automobile trips have declined; the increase was spread over trips both toward and away from the CBD. In essence, people may be driving farther to reach shopping destinations. There are, however, major changes in direction by public transit. Trips toward the CBD decreased dramatically for commuter rail, bus, and subway. Because local destinations also increased by transit (except commuter rail), the mode shift reflects the decentralization of shopping from areas in and near the CBD.

CONCLUSIONS AND IMPLICATIONS

This study offers definitions of reverse commuting and directional travel and presents a framework for their analysis. The data presented here suggest that the four definitional categories—toward the CBD, away from the CBD, local, and lateral—provide a workable framework. A substantial proportion of trips are either local or are longer but do not really make progress toward or away from the CBD. The remaining trips are clearly toward or away from the CBD, and they illustrate distinct patterns by the trip purpose and mode used.

Commuting to work has had and still exhibits a strong directional bias toward the CBD. All the discussion about reverse commuting suggests a dramatic phenomenon based on actual numbers, but the data here do not show a statistically significant increase in the proportion, especially for work trips. Approximately one-half of the work trips remain destined toward the CBD, and fewer than a fourth are away from it. However, the magnitude of reverse commuting is

TABLE 6 Direction of Work Trip by Mode for 1970 and 1990 (Unlinked Trip Direction)

Trip Mode	Toward CBD		Away from CBD		Local		Lateral	
Year	1970	1990	1970	1990	1970	1990	1970	1990
Walk	2%	3%	2%	1%	92%*	92%*	4%	5%
Auto Drive	47%	45%*	26%	28%	14%	16%	14%	12%
Passenger	42%	37%*	27%	26%	16%	23%	16%	14%
Suburban Bus	NA**	35%	NA	37%*	NA	16%	NA	12%
Commuter RR	96%*	96%*	3%	3%	0%	0%	0%	0%
City Bus	56%*	57%*	19%	17%	10%	12%	15%	14%
Subway/El	85%*	80%*	11%	14%	1%	2%	3%	4%
Taxi	22%	39%	13%	7%	49%*	40%*	16%	14%
Other	48%*	30%	19%	20%	21%	38%*	12%	12%
All Trips	49%	46%	22%	23%	17%	20%	12%	11%

* Highest values for trip mode.
** In 1970 suburban and city bus trips were combined.
Source: Computed from the 1970 and 1990 CATS Household Travel Surveys.

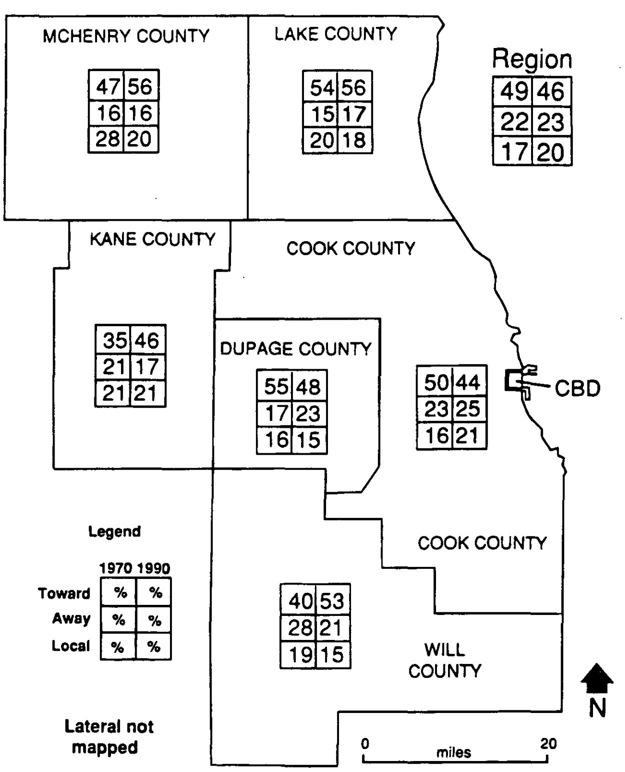


FIGURE 5 Unlinked work trip directions for 1970 and 1990.

growing because the labor force has increased, especially because the city of Chicago's population declined during the study period and few commuters use public transit in the reverse direction. Reverse commuting congestion would be worse if the proportion of the labor force commuting in this direction also increased.

In many communities reverse travel is increasing, and a substantial portion of this activity is not easily accomplished on public transit. Nevertheless, there is evidence that transit is serving some reverse travel. Table 3 shows that suburban bus was used principally to travel in the reverse direction.

The data also indicate that the reverse commute is not strictly a city-to-suburb phenomenon. It is almost as prevalent in the inner suburbs. The reverse commute drops from only 27 percent of all trips in the city to 23 percent in suburban Cook and 23 percent in DuPage counties. Considering just automobile drivers, reverse commuting is more evident, with 39 percent commuting away from the CBD and 37 percent toward it from city origins.

Regarding home-based directions it may be surprising that all trip purposes exhibit at least some bias toward the CBD as opposed to away from it. However, some of these purposes, such as shopping and recreation, do not show strong differences and as such do not substantially contribute to inbound traffic activity. The mode that shows the greatest CBD directional bias is walking. Apparently those who make trips toward the urban periphery walk very little. The bulk of the walking activity in the Chicago area, outside the local neighborhood, occurs toward the CBD.

In conclusion, besides being able to analyze the directional patterns associated with trip making, the methods used here are trans-

TABLE 7 Direction of Shopping Trips by Mode for 1970 and 1990 (Unlinked Trip Direction)

Trip Mode	Toward CBD		Away from CBD		Local		Lateral	
Year	1970	1990	1970	1990	1970	1990	1970	1990
Walk	0%	3%	0%	1%	100%	94%*	0%	2%
Auto Drive	20%	22%	21%	25%	41%*	36%*	17%	17%
Passenger	23%	22%	24%	29%	33%*	32%*	20%	17%
Suburban Bus	NA**	25%	NA	46%*	NA	15%	NA	14%
Commuter RR	38%	29%	62%*	71%*	0%	0%	0%	0%
City Bus	41%*	20%	20%	33%*	23%	31%	16%	16%
Subway/El	77%*	46%*	18%	43%	1%	11%	4%	0%
Taxi	35%	-%	7%	68%*	48%*	16%*	10%	17%
Other	0%	19%	20%	23%	81%*	45%*	0%	14%
All Trips	23%	20%	22%	24%	38%	40%	18%	15%

* Highest values for trip mode.
** In 1970 suburban and city bus trips were combined.
Source: Computed from the 1970 and 1990 CATS Household Travel Surveys.

ferable to any region with geographically coded trip data. In this way comparisons between regions can provide a better understanding of urban form and its travel implications. Finally, the directional definitions proved useful in this study, and others are invited to use them.

REFERENCES

1. Gillard, Q., Reverse Commuting and the Inner City Problem. *Growth and Change*, Vol. 10, July 1979, pp. 12–18.
2. Farkas, A. Reverse Commuting: Prospects for Job Accessibility and Energy Conservation. In *Transportation Research Record 1349*, TRB, National Research Council, Washington, D.C., 1992, pp. 85–92.
3. *A Study of Reverse Commuters in the Chicago Metropolitan Area.* Total Research Corporation, Princeton, N.J., Dec. 1991, p. 75.
4. Jarzab, J. A Note on Reverse Commuting. *Operations Review*, April 1985, pp. 5–6.
5. Marston, J., J. Walsh, K. Wertheimer, and L. Yang. Complex Trip Chains and Their Effects on Transit Ridership. In *Proc., Metropolitan Conference for Public Transportation Research*, Vol. 9, 1994.
6. Giuliano, G. Is Jobs-Housing Balance a Transportation Issue? In *Transportation Research Record 1305*, TRB, National Research Council, Washington, D.C., 1991, pp. 305–312.
7. Giuliano, G., and K. Small. Is the Journey to Work Explained by Urban Structure? *Urban Studies*, Vol. 30, 1993, pp. 1485–1500.
8. Miron, J. Centrifugal Relocation and Back-Commuting from the Metropolitan Fringe. *Journal of Transport Economics and Policy*, Vol. 16, 1982, pp. 239–257.
9. Sööt, S. Transportation Costs and Urban Rent Theory. *Land Economics*, May 1974, pp. 193–195.
10. *Final Report: Chicago Area Transportation Study 1970 Travel Surveys.* Creighton Hamburg, Inc., Chicago, Ill., July 1971.
11. *CATS 1990 Household Travel Survey: A Methodological Overview.* Working Paper 94-05. Chicago Area Transportation Study, Chicago, Ill., April 1994.
12. Lave, C. *Things Won't Get a Lot Worse: The Future of U.S. Traffic Congestion.* Institute of Transportation Studies, University of California at Irvine, 1990.
13. *Transportation Facts: A Focus on Vehicle Registrations.* Vol. 10, No. 7, Dec. 1993.
14. *Transportation System Development Plan and Transportation Improvement Conformity Documentation for Northeastern Illinois.* Chicago Area Transportation Study, Chicago, Ill. June 1994.
15. Newman, P., and J. Kenworthy. *Cities and Automobile Dependence: An International Sourcebook.* Glower, Aldershot, U.K., 1989.
16. Levinson, D., and A. Kumar. The Rational Locator, Why Travel Times Have Remained Stable. *American Planning Association Journal*, Summer 1994, pp. 319–332.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Analysis and Use of 1990 Census Transportation Planning Package in Delaware Valley Region

THABET ZAKARIA

The 1990 Census Transportation Planning Package (CTPP) for the Delaware Valley Region is analyzed, with special emphasis on journey-to-work trips, employment, mode of transportation to work, travel time, vehicle ownership, employed persons, and other socioeconomic data essential to transportation planning and travel forecasting. A review of the CTPP computer tapes and data showed some programming, sampling, and bias problems, which were resolved before the data were used as a base for trend analysis, traffic simulation, highway and transit project studies, strategic planning, and economic development. The CTPP information should be adjusted before it is used for transportation planning. The errors in the 1990 CTPP data are generally small, but the package shows no improvement over the 1980 data. Most of the 1990 CTPP problems can be avoided in the future if the recommendations made in this paper are considered in the 2000 census.

INTRODUCTION

Information on 1990 census work trips, employed persons, employment, and many other socioeconomic variables is available in the 1990 Census Urban Transportation Planning Package (CTPP). The CTPP is a special tabulation of census data used in transportation planning at the state and regional levels. The tabulations and data items were specified by an ad hoc committee of transportation planners representing TRB's Committee on Transportation Information Systems and Data Requirements. Funding for the development and production of the CTPP was provided by the states through AASHTO.

On June 22, 1993, the Delaware Valley Regional Planning Commission (DVRPC) received the first three parts of the CTPP Statewide Element, but the first three parts of the Urban Element were not received until April 21, 1994, more than 4 years after Census Day in 1990. Work was initiated to process and print CTPP data for various levels of geographic units for purposes of transportation system planning analysis and evaluation and for project studies. Because the contents of the CTPP are extensive, work on the processing and evaluation of data is still under way and is continuing in 1995.

The purpose of this paper is to discuss briefly the experience of DVRPC with the CTPP data, with special emphasis on the journey-to-work information and other socioeconomic information required for transportation planning, such as population, households, employed persons, vehicle availability, and employment. Some specific problems found in the CTPP information are defined, and some solutions are suggested. The data are evaluated and some figures are presented to illustrate the magnitude of the errors and discrepancies in the data selected. The use of CTPP data in several DVRPC transportation and nontransportation planning projects is described.

The DVRPC region includes four suburban counties in Pennsylvania (Bucks, Chester, Delaware, and Montgomery), four suburban counties in New Jersey (Burlington, Camden, Gloucester, and Mercer), and the city of Philadelphia. The Delaware Valley includes an area of 9,886 km^2 (3,817 m^2) and a population of approximately 5.2 million. There are 352 municipalities, including such major cities as Trenton and Camden in New Jersey and Philadelphia and Chester in Pennsylvania (Figure 1).

Essentially, this paper is an update of a similar paper published by the author in 1984 on the 1980 Urban Transportation Planning Package (UTPP) (*1*).

CONTENTS OF 1990 CTPP

The CTPP information was selected from the responses to the 1990 long-form census questionnaire distributed to approximately 17 percent (one in six) of all households. The Bureau of the Census prepared two 1990 CTPP packages—Statewide Element and Urban Element. The Statewide Element consists of six parts, which contain information at the municipal level [Minor Civil Divisions (MCDs)]. These parts are labeled A through F:

A. Worker and household characteristics by place of residence,
B. Worker characteristics by place of work,
C. Worker characteristics by place of residence to place of work,
D. Worker and household characteristics by large place (75,000+ population) of residence,
E. Worker characteristics by large place (75,000+ population) of work, and
F. Worker characteristics by large place of residence to large place of work.

The Urban Element provides data at the traffic analysis zone (TAZ) level. There are 1,395 TAZs in the DVRPC region, which are for the most part equivalent to census tracts. However, census block groups are used in densely developed areas, such as the Philadelphia central business district, where census tracts are too large for traffic simulation and analysis. There are eight parts in the CTPP Urban Element, labeled 1 through 8 (part 5, however, has been eliminated):

Delaware Valley Regional Planning Commission, Bourse Building, 111 South Independence Mall East, Philadelphia, Pa. 19106-2515.

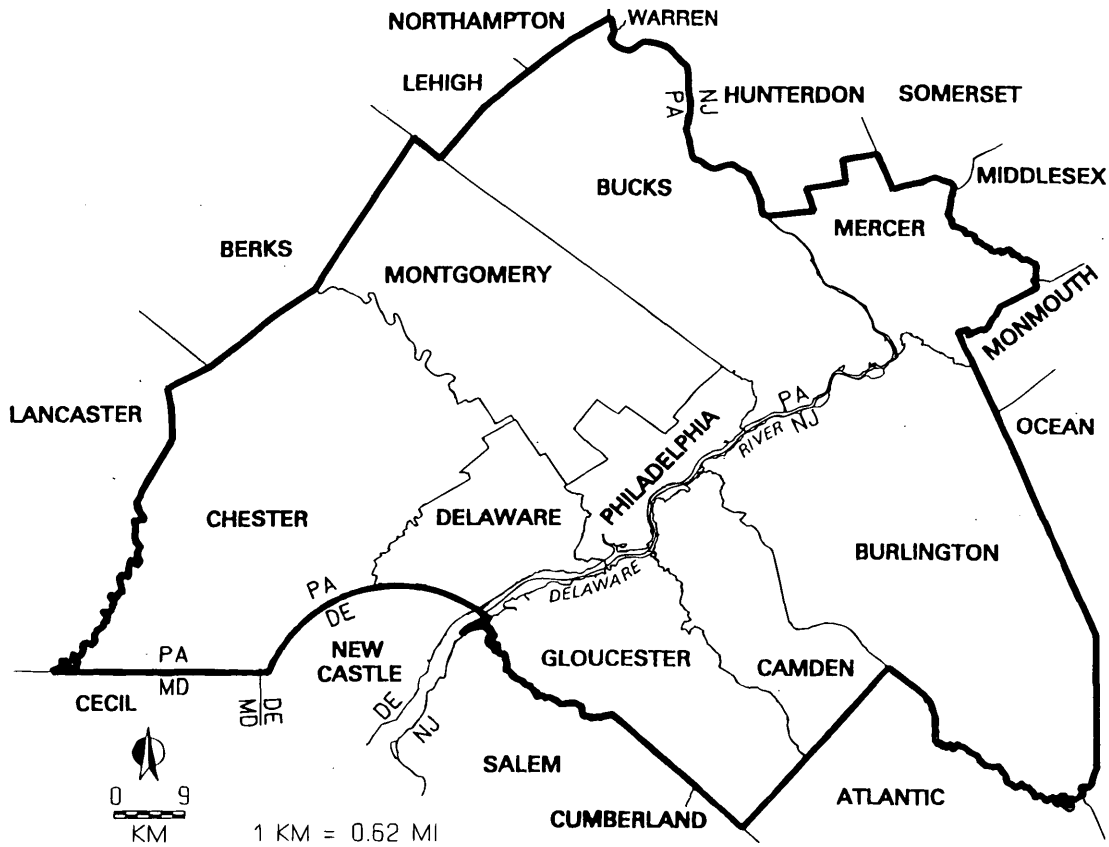


FIGURE 1 Map of the Delaware Valley region.

1. Worker and household characteristics by zone of residence,
2. Worker characteristics by zone of work,
3. Worker characteristics, zone of residence to zone of work,
4. Worker and household characteristics by superdistrict of residence,
5. Eliminated,
6. Worker characteristics by superdistrict of residence to superdistrict of work,
7. Worker characteristics by census tract of work, and
8. Detailed worker characteristics, zone of residence to zone of work (large regions of 1,000,000 population only).

It should be noted here that the 1980 UTPP included only six parts, as follows:

I. Workers and household characteristics by zone of residence,
II. Workers by zone of work for large geographic areas,
III. Data by zone of work at the tract level,
IV. Tabulations between zone of residence and zone of work at the tract level,
V. Tabulations of the zone of work data at the block-group level aggregated to census tract; and
VI. Data between place of residence and place of work at the county level, including 20 counties external to the DVRPC region that have a significant flow of work trips to and from the region.

The 1990 data were collected using census areal units consisting of block, block group, tract, enumeration district, MCD (township, borough, city, and village), county, and standard metropolitan statistical area. In 1975 the DVRPC zonal system, used for the collection of data in 1960 origin and destination surveys, was converted to the census areal system. This conversion has made it much easier to provide an equivalency table of all tracts, blocks, enumeration districts, and TAZs. The preparation of such a tabulation proved to be tedious, costly, and time consuming, as the Delaware Valley Region includes more than 74,000 blocks, 1,317 tracts, and 1,395 TAZs.

In March 1992, DVRPC prepared a correspondence table for use in the tabulation of the Urban Element information. This table includes the following:

■ TAZ, ■ Superdistrict, ■ Census state code, ■ Census tract number, ■ Census block group number, ■ MCD, and ■ County.

In addition to this information, DVRPC specified all external counties that have significant community flow from and to the DVRPC region. Because of programming difficulties, this file was not used by the Census Bureau. Instead, DVRPC was asked in 1994 to prepare a revised correspondence table (equivalency file), which inserted the TAZs in each census block record.

ANALYSIS AND EVALUATION OF 1990 CTPP DATA

A review of the 1990 CTPP data of parts 1–3 for the Delaware Valley region indicated some programming, definitional, and statistical problems. Unlike those for the 1980 UTPP, the 1990 data on work-trip destinations contain trips not identified by block or tract. The

Census Bureau could not allocate all 1990 trips to TAZs, because the Topologically Integrated Geographic Encoding and Referencing (TIGER) file does not contain address ranges for some suburban and rural areas in the region. The Census Bureau provided a list of places that filed the census allocation process. Specifically, any place that has less than 70 percent address range coverage and less than 70 percent of the persons working in them coded to tract and block failed the test. For such places the Census Bureau allocated the work places to default zones and asked DVRPC to review the list and allocate the default data to the affected TAZs, including water tracts (Figure 2).

Programming and Format Problems

After receiving the 1990 CTPP tapes from the Census Bureau, DVRPC immediately started to extract the data needed for various air-quality and transportation planning studies. It was found that the format of the tapes is quite complex and unclear. There was no labeling on the tapes, and the names of the tables were confusing. No documentation of certain record types was available. The variations in recorded content should have been clearly documented both in the general documentation and in the data dictionary.

For example, review of the tapes of the CTPP Urban Element indicated that they do not have the same computer record size and block size at the tract and zonal levels. Part 3 has a record size of 1,180 and block size of 23,600, but Part 2 has a record size of 10,616 and block size of 21,230. These problems have caused some delay and duplication of effort.

Problems of Definition and Statistics

As stated previously, the Census Bureau obtained information on workers and not on trips; the latter information is usually collected in home interview surveys for transportation planning studies. The questions about mode of transportation in the census for 1980 and 1990 are similar, as can be seen below:

1980	*1990*
Car	—
Truck	—
Van	Car, truck, or van
Bus, streetcar	Bus or trolley bus
—	Streetcar or trolley car
Subway, elevated	Subway or elevated
Railroad	Railroad
—	Ferry boat
Taxicab	Taxicab
Motorcycle	Motorcycle
Bicycle	Bicycle
Walked only	Walked
Worked at home	Worked at home
Other	Other

The analysis of workers' trip tables (Part 3) by travel mode indicated that some walking and railroad trips were unrealistic in terms of travel time or distance. It was found, for example, that some workers walked from Philadelphia to places a considerable distance from the city. Similarly, there were railroad trips for which no such service existed. These few irrational trips are due to errors in census coding, to sampling error, or to incorrect information returned by respondents who did not understand the census questionnaire. Many respondents confused the access mode with the principal mode of travel.

The evaluation of employment data by industry showed that some respondents misunderstood the census question that used the Standard Industrial Classification (SIC) system (Question 28). Some were not able to identify their industry correctly because some SIC categories are not easily defined. The public administration sector is especially complicated. Employees of a municipal utility authority, for example, may consider themselves members either of the public administration sector or of the public utilities sector.

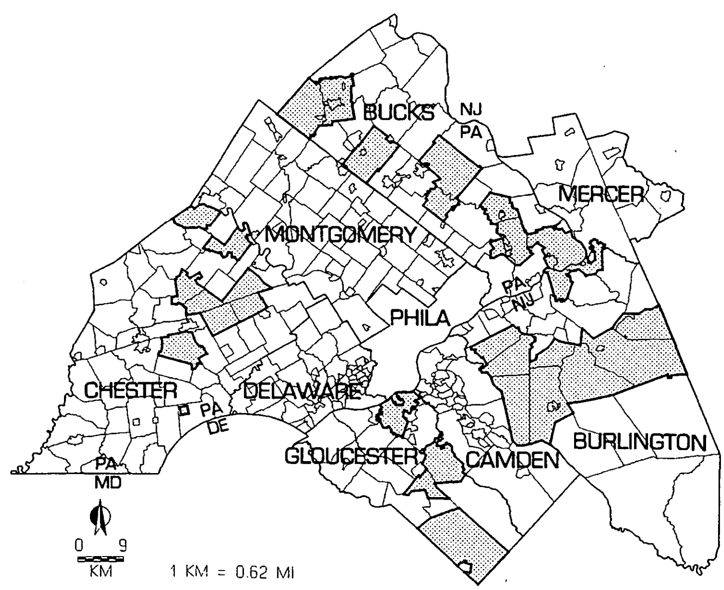


FIGURE 2 Location of default TAZs.

Accuracy of CTPP Data

Generally, the 1990 CTPP data are good for transportation planning purposes. The data on population, household, car ownership, employed persons, and other socioeconomic characteristics obtained from part 1 are quite accurate and do not require any major adjustment for sampling or nonsampling errors.

Part 1 data compare favorably with the 100 percent census counts. Table 1 shows the magnitude of difference between the population produced from Part 1 and from the 100 percent counts for a few TAZs, MCDs, and counties selected at random. As can be seen, the differences in population and resident workers are small and are acceptable for all planning purposes. However, the number of households in the CTPP is slightly lower than the total counts. Most of the difference, 2,319 of a total of 2,667, is found in the city of Philadelphia. All household zonal data were adjusted to be consistent with the 100 percent census counts, which are equal to those extracted from the tabulations on Standard Tape File 3.

As described previously, Parts 2 and 3 contain trip data at the place of work for various geographic units such as TAZs, MCDs, and counties. If trip destinations by resident and nonresident workers living in commutershed areas are summed, the total would be approximately equal to the number of jobs, or employment. A certain percentage of these work-trip destinations (employment) should be added to account for workers who were absent during the census week because of illness or vacation or for other personal reasons and for workers who had more than one job. Data from the Bureau of Economic Analysis, the Bureau of the Census, the Bureau of Labor Statistics, previous DVRPC employment files, and local wage records, were used to adjust the CTPP work trips (number of workers at the place of destination) four times to develop the employment file. The first adjustment was made to account for absentee rates reported by the census for each county (2.16 percent for the region) from responses to Question 21a on the long form used in 1990.

Second, all employment data were adjusted upward to reflect multiple job holding, using information from a survey conducted for the Bureau of Labor Statistics with a national sample of approximately 60,000 households (*2*). It was found that the national rate for multiple job holding is 6.2 percent and varies by employment sector, ranging from 4.7 percent for construction workers to 9.3 percent for those working in government. Third, employment estimates at the municipal level were examined, and some were adjusted upward or downward to account for coding discrepancies and respondent errors. Such adjustments were necessary at the municipal level to bring the estimates into agreement with data from the Bureau of Labor Statistics, DVRPC files, and municipal tax records.

Finally, employment estimates at the TAZ level were examined to allocate the trips coded to default zones and water tracts. All zonal data were factored to county and municipal control totals by employment sector, and a new computer file was prepared for users of these data. DVRPC uses the following 11 SIC sectors in the travel simulation process: agriculture, forestry, and fisheries; mining; construction; manufacturing; transportation; wholesale trade; retail trade; finance, insurance, and real estate; service; government; and military.

Table 2 shows a comparison of CTPP employment before and after adjustments for selected TAZs, MCDs, and counties and for the total region. It also shows the percent difference between the unadjusted CTPP employment estimates and those adopted by DVRPC. As shown in Table 2, the differences between the two sets of employment estimates are small (approximately 10 percent). In general, because of sampling error and coding problems, the percent difference between the two sets of employment estimates increases as the size of a geographic unit decreases.

Most parts of the CTPP include information on the worker's mode of transportation to work. A respondent was asked to choose one of eleven travel modes that he or she usually took to work for most of the distance between the place of residence and work. The travel mode proportions appear to be reasonable because they compare favorably with DVRPC highway traffic counts and transit surveys for large areas and the region. Table 3 shows that the difference between the CTPP data and DVRPC estimates for total public transportation work trips is 1.5 percent. However, such a difference

TABLE 1 Comparison of 1990 CTPP Population, Households, and Resident Workers with Total Census Counts

	Population			Households			Resident Workers		
Areal Unit	CTPP	Total Count	Diff.	CTPP	Total Count	Diff.	CTPP	Total Count	Diff.
TAZ									
100	1,204	1,205	−1	476	515	−39	557	558	−1
400	9,030	9,030	0	4,279	4,319	−40	3,678	3,678	0
700	3,750	3,750	0	1,459	1,434	25	2,118	2,118	0
990	6,169	6,169	0	1,880	1,881	−1	2,780	2,780	0
Municipality									
New Hope, PA	1,400	1,400	0	811	810	1	964	964	0
Media, PA	5,957	5,957	0	2,876	2,867	9	3,243	3,243	0
Glassboro, NJ	15,614	15,614	0	5,069	5,019	50	7,422	7,422	0
County									
Philadelphia, PA	1,585,577	1,585,577	0	600,740	603,059	−2319	657,389	657,387	2
Mercer, NJ	325,824	325,824	0	116,777	116,941	−164	166,680	166,680	0
Total Region	5,182,705	5,182,705	0	1,891,614	1,894,281	−2667	2,496,292	2,496,215	77

TABLE 2 Comparison of 1990 CTPP and DVRPC Adopted Employment Estimates

Areal Unit	1990 Employment Estimates CTPP Unadjusted	CTPP Adjusted	DVRPC Adopted	% Diff. Adopted vs. CTPP Unadj.
TAZ				
100	119	119	–	–
400	1,711	1,719	–	–
700	2,259	2,349	–	–
990	492	493	–	–
Municipality				
New Hope, PA	2,351	2,579	2,351	0.0
Media, PA	10,110	10,993	11,210	10.9
Glassboro, NJ	7,287	7,924	7,924	8.7
County				
Philadelphia, PA	761,244	834,335	836,874	9.9
Mercer, NJ	204,826	224,356	220,592	7.7
Total Region	2,433,682	2,697,229	2,693,879	10.7

TABLE 3 Comparison of 1990 CTPP and DVRPC Work Trip Estimates by Mode (thousands of trips per day)

Areal Unit	Mode	CTPP Data	DVRPC Estimates	Percent Diff.
	Regional Rail	27.7	33.5	20.9
Philadelphia	Subway–Elevated	33.6	41.9	24.7
CBD	Surface Transit	47.9	62.9	31.3
	Highway	117.9	74.8	−36.6
Total Region	Public Transportation	273.2	277.2	1.5
	Highway	1,954.0	1,792.0	−8.3

becomes large for travel submodes within smaller areas. In the Philadelphia central business district the difference between the CTPP and estimated subway-elevated trips is approximately 25 percent. Such large differences are due mainly to incorrect responses to the questionnaire. It appears that many respondents confused the access mode to a station with the principal mode of travel to work. For example, persons who live in Delaware County and work in the Philadelphia center city must take buses or trolleys to the 69th Street Terminal, where they transfer to the Market-Frankford subway-elevated line. Thus, some respondents reported bus or trolley as the principal means of transportation rather than subway, which was the correct response. As shown in Table 3, the surface trips (bus and trolley) are overestimated by 31.1 percent. The highway trips estimated by the DVRPC model are underestimated. This problem will be resolved when the model is calibrated with the 1990 census data.

Table 4 shows the 1980 and 1990 average travel time of commuters in selected counties and total region. The regional average travel time has changed slightly since 1980. The changes observed in travel times at the county level are also very small.

Despite the increasing traffic congestion in the region, the regional travel time of work trips declined by 2.8 percent in the 1980s. This is due to the decline of urban areas and to growth in the suburbs, where the private automobile is the predominant mode of travel. According to the CTPP data, it takes much longer to commute by public transportation than by automobile. The decline of the share of public transportation in the region contributes to the decrease of commuting times because driving alone takes less time. Commuters in the region have shifted from slower to faster modes of transportation. The 1990 CTPP average regional travel time compares very well with DVRPC average community time, which is based on actual travel survey (24.6 versus 22.8 min).

These problems are similar to those experienced with the 1980 UTPP. For this reason the CTPP trip information should be adjusted before it is used for transportation planning. The adjusted CTPP employment and traffic data for the Delaware Valley region are quite reasonable.

TABLE 4 Average Travel Time of Commuters

County of Residence	Travel Time (minutes) 1980	1990	% Diff
Bucks, PA	24.0	24.2	0.8
Montgomery, PA	21.9	22.5	2.7
Burlington, NJ	24.0	24.1	0.4
Mercer, NJ	21.7	22.1	1.8
Total Region	25.3	24.6	−2.8

USES OF 1990 CTPP AT DVRPC

The uses of the 1990 CTPP in the Delaware Valley region are somewhat similar to applications in other metropolitan areas (*3–6*). DVRPC has already used census data in various transportation planning studies and will continue to use the CTPP in the future. As mentioned earlier, the CTPP includes many socioeconomic data items and trip information that are invaluable to local and state governments, transit operators, and private corporations for making a variety of transportation and locational decisions. These include such decisions as the locations of shopping centers, industrial parks, banks, and service industries and the estimation of highway and transit travel, parking requirements, transit fleet sizes, and service schedules.

There are at least six major uses for the 1990 CTPP in the Delaware Valley region. Some of these have been applied and some will continue in the future.

Development of Data Base for Transportation Planning

DVRPC has initiated a project to prepare a data bank for transportation planning at the TAZ, superdistrict, and municipal levels. This information includes population, vehicle availability, employment, work trips by mode, travel time, household income, and other socioeconomic variables required for traffic simulation and transportation analysis and planning. Such data have been extracted from Parts 1–3 of the CTPP. All data items have been edited for reasonableness and adjusted if necessary on the basis of other census data and DVRPC surveys, counts, and data as described in the previous section of this paper. These data will be used in most transportation system and project planning studies in the next 10 years.

Preparation of Data Summaries and Evaluation of Trends

DVRPC has completed three reports on the journey-to-work trends in the Delaware Valley region (*7–9*). These reports compare the 1970, 1980, and 1990 journey-to-work information, means of transportation for commuting to work, employed persons, and employment at the county and regional levels. They also analyze the commuting flow between the counties of the Delaware Valley region and surrounding counties and cities. The reports were well received by planners and decision makers because they provide factual information about trends in development and travel patterns in the region. For example, Table 5, taken from the regional report (*7*), gives the 1970–1990 trends in the distribution of Montgomery County resident workers by place of work. Other tables show the trends in employment and mode of travel for all DVRPC counties, cities, and selected municipalities.

Short data bulletins were also published. Each includes one or two information items obtained from Part 1, 2, or 3 of the CTPP. For example, a bulletin was prepared on vehicle ownership growth between 1970 and 1990 for the counties in the Delaware Valley region. It also includes households stratified by the number of vehicles owned (zero, one, two, or three or more cars).

Update of DVRPC Traffic Simulation Models

A project has been initiated to update the DVRPC travel forecasting models by using the 1990 CTPP. During the 1980s the 1980 UTPP was used to check and validate the DVRPC traffic simulation models (Figure 3). These models will be updated again using 1990 census data. The DVRPC travel simulation models follow the tradi-

TABLE 5 Montgomery County Resident Workers Distribution by Place of Work

Place of Work	1970	1980	1990	Percent Change '70–'80	Percent Change '80–'90
Bucks	8,488	14,325	20,986	68.8	46.5
Chester	5,900	10,525	17,920	78.4	70.3
Delaware	5,897	7,773	10,933	31.8	40.7
Montgomery	158,986	204,673	229,923	28.7	12.3
Philadelphia	54,489	55,598	55,956	2.0	0.6
Total PA portion	233,760	292,894	335,718	25.3	14.6
Burlington	1,632	532	1,484	−67.4	178.9
Camden	3,089	1,643	2,808	−46.8	70.9
Gloucester	883	225	474	−74.5	110.7
Mercer	1,877	354	1,024	−81.1	189.3
Total NJ portion	7,481	2,754	5,790	−63.2	110.2
Total Region	241,241	295,648	341,508	22.6	15.5
Berks	2,499	3,070	3,670	22.8	19.5
Lancaster	82	172	162	109.8	−5.8
Lehigh	633	773	1,390	22.1	79.8
New Castle	513	282	580	−45.0	105.7
Northampton	665	196	326	−70.5	66.3
Other	5,504	4,185	5,324	−24.0	27.2
Total External	9,896	8,678	11,452	−12.3	32.0
Total Trips	251,137	304,326	352,960	21.2	16.0

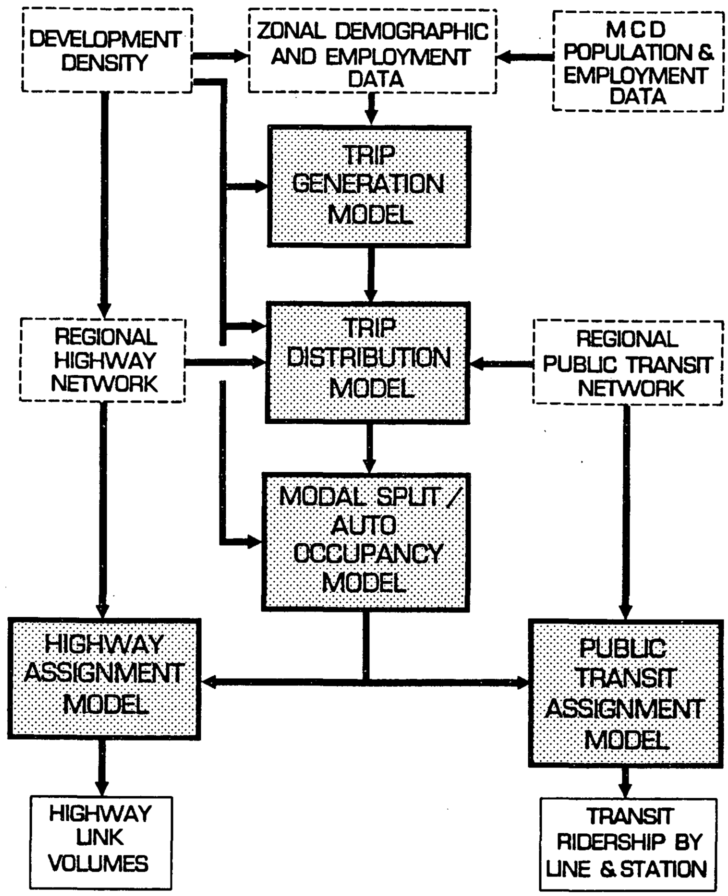


FIGURE 3 DVRPC regional travel simulation process.

tional steps of trip generation, trip distribution, modal split, and travel assignment and use the computer programs included in the Urban Transportation Planning System. In addition to this system, DVRPC is now using the TRANPLAN and TRANSCAD systems for travel forecasting and air-quality analysis.

Generally, the models are similar to those used in other large urban areas that depend on census data for system and project studies. Figure 3 shows the steps needed to update the DVRPC traffic simulation process. This work will be completed by the end of 1995. A careful review and evaluation of the results of each model will be conducted, and necessary adjustments will be made to achieve the most accurate calibration. The simulated traffic volumes will be compared with actual highway traffic counts and public transportation ridership to ensure that acceptable accuracy of the simulated results is obtained from these models. Specifically, DVRPC will use the CTPP data in the following activities:

- Development of accurate inputs on population, households, vehicle availability, resident workers, and employment at the TAZ level;
- Comparison and analysis of DVRPC trip rates for work with CTPP;
- Comparison and analysis of DVRPC trip length and travel time distribution for work with CTPP;
- Comparison and evaluation of work trips estimated by the DVRPC model with CTPP;
- Comparison and analysis of DVRPC auto occupancy model with CTPP; and
- Analysis and evaluation of DVRPC external work trips with CTPP.

Use in Highway and Transit Corridor Studies

The 1990 CTPP data, especially the journey-to-work information contained in Part 3, have been used in several transit corridor studies to check the travel demand or ridership for each transit submode, including high-speed rail line, express bus and park-and-ride service, and local bus service in the suburbs.

The 1990 CTPP data will also be used in many future highway and transit corridor studies, because it is the only information available for transportation planning at the regional level. The use of the CTPP minimizes any large-scale data collection in the Delaware Valley and decreases the rising costs of surveys required for transportation planning at the system and project levels.

Application in Strategic Planning and Economic Development

DVRPC is planning to use the 1990 CTPP information on employment, particularly Part 2, to evaluate the significant changes in the type and location of industries and commercial establishments. This evaluation will result in recommendations and strategies aimed at attracting new industries and high-technology firms to the Delaware Valley. Also, employment information is useful to the redevelopment of declining areas of old urban centers and provision of the required physical improvements for their rehabilitation.

Based on the inventory of major employment centers, a project has been developed to analyze these centers' accessibility for workers. Existing access patterns will be examined in terms of origin–destination and model split. Access problems for particular labor populations such as low-wage workers will be studied. Access problems and opportunities will be identified based on data analyses, site analyses, and consultations. Both transportation and land-use-related solutions will be proposed to improve access to employment centers, as dictated by findings.

Provision of 1990 CTPP Data to Public Agencies and Private Corporations

Finally, DVRPC intends to provide the 1990 CTPP information to any public or private agency involved in planning or urban studies such as studies for housing, finance, real estate, health facilities, social services, economic base, and economic development. It appears that many planning agencies and private companies in the Delaware Valley region are very much interested in obtaining the CTPP information.

CONCLUSIONS

Generally, the 1990 CTPP for the Delaware Valley region contains valuable data for air-quality and transportation planning, economic base and employment location studies, urban development analysis, and planning and evaluation of transit services. However, the analysis of CTPP data indicates a few programming, statistical, and bias problems. Most of these problems were resolved before DVRPC used the CTPP as a data base for trend analysis, information purposes, traffic simulation, highway and transit project studies, strategic planning, and economic development. The errors in the 1990 data are generally similar to those found in the 1980 UTPP. Like the 1980 data, the 1990 employment estimates must be adjusted before they are used in transportation planning studies because they do not include all workers or jobs.

Most of the 1990 CTPP problems and errors can be avoided in the 2000 census by quality-control edits and a careful review of the census questionnaire as well as of the computer formats and programs required for processing the information. Specifically, the journey-to-work questions (22 and 23) should be simplified to prevent any confusion on the part of respondents on such questions as mode of travel and industry classification. Many confused the access mode to subway-elevated or railroad lines with the principal mode of travel. The questionnaire should be redesigned to capture multimodal trip information from the place of residence to the place of work. Question 28 should be simplified to avoid any error or misunderstanding in the employment sectors.

The format of the 1990 CTPP tapes is rather complex and must be simplified and checked for consistency. The funding and development of two packages in 1990—State and Urban Elements—is an excellent idea, because these packages include better coverage of commutershed areas and could be used for checking and consistency of the census information. AASHTO should again provide the funding for the 2000 CTPP. Finally, DVRPC has not as yet received all parts of the CTPP; a more timely release of data is obviously important to all census data users.

ACKNOWLEDGMENT

This paper was financed in part by FHWA and FTA, of the U.S. Department of Transportation, and the Pennsylvania and New Jersey departments of transportation.

REFERENCES

1. Zakaria, T., Analysis and Use of 1980 Urban Transportation Planning Package in the Delaware Valley Region. In *Transportation Research Record 981*, TRB, National Research Council, Washington, D.C., 1984, pp. 95–102.
2. *Multi-Job Holding Unchanged in May 1991.* Report 91-547. Bureau of Labor Statistics, Oct. 1991.
3. Goodman, C., and E. Murakami. Using 1990 Census Data in National Policy Analysis. In *Conference Proceedings 4: Decennial Census Data for Transportation Planning*, TRB, National Research Council, Washington, D.C., 1995, pp. 105–114.
4. Purvis, C. The Decennial Census and Transportation Planning: Planning for Large Metropolitan Areas. In *Conference Proceedings 4: Decennial Census Data for Transportation Planning*, TRB, National Research Council, Washington, D.C., 1995, pp. 55–67.
5. Cervero, R. Use of Census Data for Transit, Multimodal, and Small-Area Analyses. In *Conference Proceedings 4: Decennial Census Data for Transportation Planning*, TRB, National Research Council, Washington, D.C., 1995, pp. 83–94.
6. Rossetti, M. A., and B. S. Enersale. *Journey-to-Work Trends in the United States and Its Major Metropolitan Areas, 1960–1990.* FHWA, U.S. Department of Transportation, Nov. 1993.
7. *Journey-to-Work Trends in the Delaware Valley Region, 1970–1990.* Direction 2020 Report 5. Delaware Valley Regional Planning Commission, Philadelphia, Pa., June 1993.
8. *Journey-to-Work Trends in Camden, Trenton, Chester, and Philadelphia, 1970–1990.* Direction 2020 Report 16. Delaware Valley Regional Planning Commission, Philadelphia, Pa., March 1994.
9. *Journey-to-Work Trends in Eight Suburban Townships, 1970–1990.* Direction 2020 Report 17. Delaware Valley Regional Planning Commission, Philadelphia, Pa. March 1994.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Estimation of Pass-By Trips Using a License Plate Survey

SOUMYA S. DEY AND JON D. FRICKER

A significant portion of the travel attracted by generators such as shopping centers and several other convenience-oriented land uses are pass-by trips. Methods exist for handling pass-by trips in traffic impact analysis. However, there is a dearth of data regarding the percentage of pass-by trips for a particular type and intensity of land use. The traditional way of determining the number of pass-by trips is to conduct a face-to-face interview survey. This is, however, a time-consuming and intrusive process. A study was conducted to determine whether it is possible to estimate the percentage of pass-by trips by using a license plate survey instead of the traditional interview survey. Two separate data collection efforts were conducted. At each location the license plate observations were matched by using a standard computer program. In both cases the percentage of pass-by trips obtained by a license plate survey was found to be very close to the results obtained by an interview survey. However, further studies must be conducted before the methodology can be accepted as a standard procedure for estimating pass-by trips. At a minimum, the procedure can place upper and lower bounds on the percentage of pass-by trips for a particular existing development. This nonintrusive method will be especially helpful (in terms of management of time and personnel) in determining the percentage of pass-by trips for large developments when a face-to-face interview would involve an extensive effort.

INTRODUCTION

Shopping centers and several other convenience-oriented land use types such as banks, gasoline stations, and fast-food restaurants have slightly different trip generation characteristics from other land use types. A significant portion of the trips attracted by these generators are "captured" from the adjacent traffic stream. Trips to such developments may be broken down into three categories (*1*): primary, diverted, and pass-by.

A primary trip destined to a retail facility is one in which the purpose of the trip is shopping at the site and the pattern of the trip is home-shopping-home.

A diverted linked trip to a retail facility is one in which the shopping destination is a secondary part of the primary trip and the pattern of the trip is, for example, work-shopping-home. A diverted linked trip as shown in Figure 1 involves a route diversion to reach the site (*2,3*).

A pass-by trip comes directly from the traffic stream passing the facility on the adjacent roadway system and does not require diversion from another roadway.

Estimating the percentage of pass-by trips is very important for many engineering studies, including traffic impact analysis. The percentage of pass-by trips varies with the size of the development, its geographical location, the nature of the roadway system, and certain other variables. Methods exist for handling pass-by trips in traffic impact analysis (*3*). However, there is a dearth of data regarding the percentage of pass-by trips for particular types and intensities of land use.

The traditional means of determining the percentage of pass-by trips is to conduct a face-to-face interview survey. Because this is a time-consuming and labor-intensive process, most of the traffic impact studies conducted across the nation either do not consider pass-by trips in their analysis or use the scatter plots and regression equations provided by ITE in its report *Trip Generation* (*2*). In fact, responses to the survey questionnaire (*5,6*) revealed that 63 percent of the states depend entirely on the ITE data to determine the percentage of pass-by trips. Twenty-six percent of the states try to use local pass-by percentages. In the absence of such data, they use the ITE rates. Eleven percent do not incorporate pass-by trips in their analyses. Therefore, 89 percent of the states depend either directly or indirectly on the ITE data base. Unfortunately, however, the size of the data base is quite small, and the regression equations, when available, have very low R^2 values (of the order of 0.3). Hence, the validity of the curves is not beyond question. Moreover, using these curves blindly would fail to take into account the site-specific characteristics of the development under consideration.

In the 1991 edition of the ITE report (*2*) a new methodology for estimating the percentage of pass-by and diverted linked trips was suggested based on the volume of the traffic available to produce pass-by and diverted linked trips (shown in Figure 2) multiplied by an attraction factor related to the size of the development. The following set of equations is suggested (*2*):

$$N_{pb} = p \times VOL_{pb}$$
$$N_d = p \times VOL_d$$
$$p = a_o + a_1 G$$

where

p = probability that a driver already in the traffic stream will stop at the generator ($0 \leq p \leq 1$),
VOL_{pb} = passing traffic stream volume available to produce pass-by trips,
VOL_d = traffic volumes on other streets available to produce diverted linked trips,
G = gross leasable area of development ($\times 1000$ m^2), and
a_0, a_1 = coefficients to be calibrated.

More recently, Moussavi and Gorman (*7*) have proposed a method for estimating pass-by trips by using multiple regression. The independent variables used in their regression analysis include floor area, volume-to-capacity (v/c) ratio, average daily traffic (ADT), and percentage of residential and commercial land use within a meter's radius of the site. These methods are also data

S. S. Dey, JHK & Associates, 1900 N. Beauregard St. Alexandria, Va. 22311-1811. J. D. Fricker, School of Civil Engineering, Purdue University, West Lafayette, Ind. 47907.

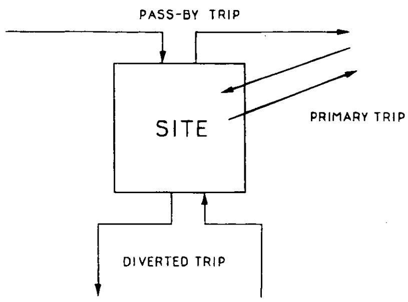


FIGURE 1 Schematic diagram showing primary, pass-by, and diverted linked trips (3).

intensive and need further evaluation. Therefore, the need for a quick and easy method to estimate pass-by trips was felt.

STUDY OBJECTIVE

The objective of the study was to determine whether it is possible to estimate the percentage of pass-by trips by using a license plate survey instead of the traditional interview survey. Two separate data collection efforts were conducted:

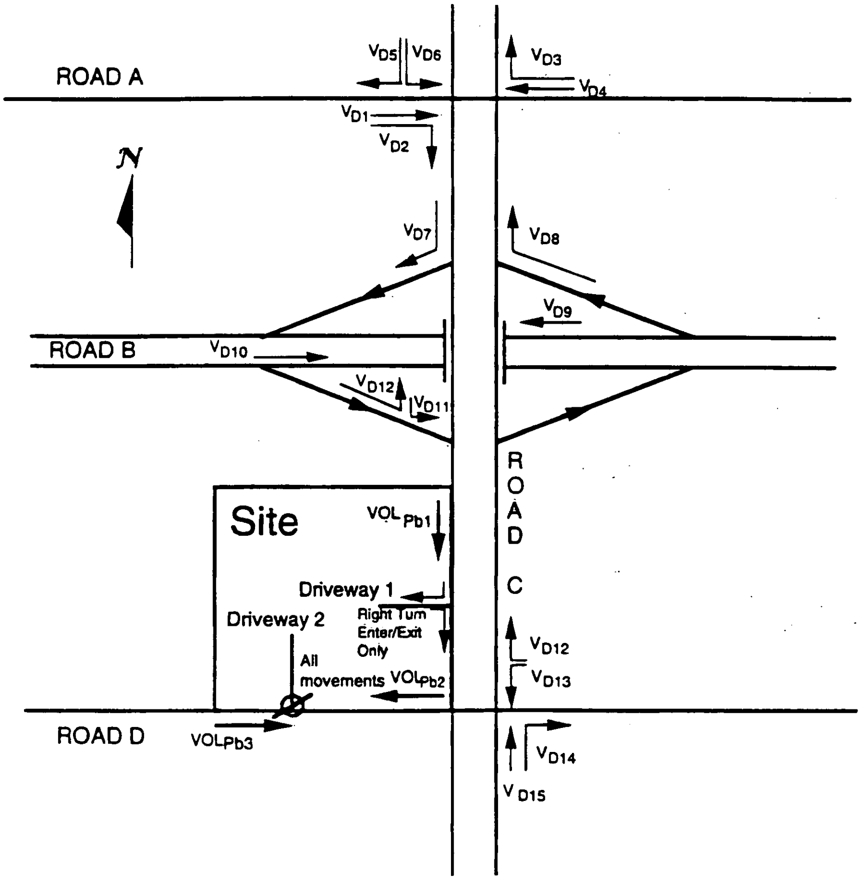
- At Eastway Plaza, a shopping center in Lafayette, Indiana; and
- At a Marsh supermarket in Castlecreek Plaza, Indianapolis, Indiana.

In both cases the percentage of pass-by trips obtained by license plate survey was compared with the results obtained by a face-to-face interview survey.

DATA COLLECTION

At Eastway Plaza the data collection was done in two parts:

1. A group of 13 interviewers intercepted people entering and leaving the stores and conducted a face-to-face interview survey.
2. One person was assigned to each of the following locations (see Figure 3): Driveway 1, access to US 52; driveway 2, access to Greenbush Street (closest to US 52), and driveway 3, secondary access to Greenbush Street.



Legend

$V_{Pb\,TOT} = \sum VOL_{Pb1} + VOL_{Pb2} + VOL_{Pb3}$ (VPH)

$V_{D\,TOT} = \sum VOL_{D1} + VOL_{D2} + VOL_{D....} + VOL_{D15}$ (VPH)

FIGURE 2 Identification of pass-by and diverted linked trip volumes (2).

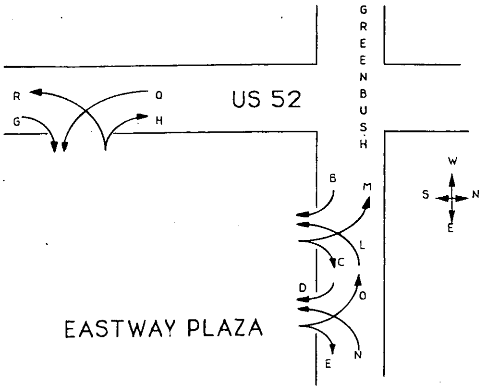


FIGURE 3 Turning movements at Eastway Plaza.

At each of these locations the last three digits of the license plates of vehicles entering and leaving the shopping center were recorded, along with the time of observation and movement of the vehicles after leaving or before entering the shopping center parking lot.

At Eastway Plaza, data were collected for 1 hr (5:00 to 6:00 p.m.) during the adjacent street evening peak period on a typical weekday.

An Indianapolis transportation engineering firm had conducted an interview survey of the customers leaving the Marsh supermarket at Castlecreek Plaza. Therefore, at this site only a license plate survey of the vehicles entering and leaving the plaza by the two driveways and of customers leaving the Marsh store was conducted. Two persons collected the license plate numbers of customers leaving the Marsh store and three persons recorded the final four license plate characters of the vehicles entering and leaving the shopping center driveways.

At Castlecreek Plaza, data were collected on a Friday for 1.5 hr, between 5:00 and 6:30 p.m.

DATA ANALYSIS

In both cases the results of the interview survey were analyzed, and the data from the license plate survey were stored in an input file.

The summary of interview survey results at Eastway Plaza is shown in Table 1. The vehicle movements at Eastway Plaza were defined as shown in Figure 3.

All possible combinations of in/out movements were identified and the license plate observations were matched using a computer program. The program identified the number of cars using each possible in/out combination. From the turning movements at the driveways and the researchers' ideas about the local travel patterns and adjacent land uses, the trips were classified as primary (PR), diverted (DI), or pass-by (PB). In many cases, however, no clear decision could be taken regarding the trip type. The results of the matches and the three predicted trip types are shown in Table 2. PB/DI and PR/DI in Table 2 denote the trips for which no decision could be made as to whether they were pass-by or diverted and primary or diverted, respectively.

The total number of matches obtained was 135. From the results of the matches, the following sets of equations were generated (see Table 3):

$$T_{PR} + T_{PB} + T_{DI} = 135 \quad (1)$$

$$T_{PB1} = 19 \quad (2)$$

$$T_{DI1} = 25 \quad (3)$$

$$T_{PB2} + T_{DI2} = 41 \quad (4)$$

$$T_{PR} + T_{DI3} = 50 \quad (5)$$

TABLE 1 Summary of Interview Survey at Eastway Plaza

STORE NAME	TOTAL # OF TRIPS	PASS-BY TRIPS	NON-PASS-BY TRIPS	NOT KNOWN
B.J.'s	5	0	5	0
Carruso's	15	2	13	0
MVP Sports	6	2	4	0
Aunt Orva's	4	1	1	2
Frame Shoppe	1	0	1	0
Queen City	5	0	3	2
Videoland	59	11	42	6
Fast Food	13	6	7	0
Radio Shack	13	5	5	3
Nutri System	10	1	5	4
Bar Barry	36	13	17	6
Progolf	4	2	2	0
Tropicana	2	0	2	0
Homework	4	0	1	3
TOTAL (PERCENTAGE)	177	43 (28.5%)	108 (71.5%)	26

TABLE 2 Results of License Plate Matches at Eastway Plaza

MOVEMENT TYPE	# OF MATCHES	TRIP TYPE	MOVEMENT TYPE	# OF MATCHES	TRIP TYPE
G→H	17	PB/DI	B→C	9	PB/DI
Q→R	3	PB/DI	B→E	10	PB/DI
N→O	0	PB/DI	N→M	1	PB/DI
L→M	1	PB/DI	L→O	1	PB/DI
D→E	1	PB/DI	D→C	0	PB/DI
G→R	17	PR/DI	Q→H	3	PR/DI
B→M	4	PR/DI	B→O	7	PR/DI
L→C	5	PR/DI	L→E	6	PR/DI
D→O	0	PR/DI	D→M	0	PR/DI
N→E	0	PR/DI	N→C	0	PR/DI
G→E	5	PB	G→O	1	DI
G→C	10	PB	G→M	12	PB/DI
Q→E	2	PB/DI	Q→O	0	PR/DI
Q→C	5	PB/DI	Q→M	2	PR/DI
N→R	0	PB/DI	N→H	0	PB/DI
D→R	0	DI	D→H	0	PR/DI
B→R	5	PB/DI	B→H	6	PR/DI
L→R	1	PB	L→H	1	PB/DI

TABLE 3 Summary of Table 2

Trip Type	Total Matches
PB	19
DI	25
PB/DI	41
PR/DI	50
Total	135

where

$$T_{PB} = T_{PB1} + T_{PB2} \quad (6)$$

$$T_{DI} = T_{DI1} + T_{DI2} + T_{DI3} \quad (7)$$

In Equations 1–7

T_{PR}, T_{PB}, T_{DI} = number of primary, pass-by, and diverted trips, respectively,
T_{PB1} = number of trips that are clearly pass-by,
T_{DI1} = number of trips that are clearly diverted,
T_{PB2}, T_{DI2} = number of pass-by and diverted trips in the PB/DI trip category, and
T_{PR}, T_{DI3} = number of primary and diverted trips in the PR/DI trip category.

T_{PB} has two components, T_{PB1} and T_{PB2}; and T_{DI} has three components, T_{DI1}, T_{DI2}, and T_{DI3}.

This set of equations reduces to

$$T_{PB2} + T_{DI2} = 41$$
$$T_{PR} + T_{DI3} = 50$$

This is a set of two equations and four unknowns that has no unique solution.

To circumvent this problem, the trip types were reduced from the three mentioned before (pass-by, diverted, and primary) to two, pass-by (PB) and non-pass-by (NPB). This was done by lumping the diverted (DI) and primary (PR) trips together into a common category, non-pass-by (NPB) trips. The equations reduced to

$$T_{PB2} + T_{NPB2} = 41$$

This is one equation and two unknowns, which also has no unique solution.

Two other ways to circumvent the problem faced are as follows:

1. Decide on the trip types based on the turning movements at the adjacent intersections. This method could yield a unique solution, but in some cases the trip type may still be debatable and subjective. For example, in the case of Eastway Plaza, if a license plate survey had been conducted at the intersections of US 52/Greenbush and the next intersection toward the south simultaneously with the surveys at the driveways, it would probably have been possible to decide on the trip types in most of the cases. Because of time constraints this method was not pursued further.
2. Split up the two trip types, pass-by and non-pass-by, and conduct an "extreme analysis." This solution would provide a range of the percentage of pass-by trips for the proposed development. From the range, a plausible percentage of pass-by trips may be estimated. The analysis was based on this approach.

The results of the matches (based on two trip types, pass-by and non-pass-by) at Eastway Plaza are shown in Table 4.

The results of the interview survey conducted by the consultants are shown in Table 5. The turning movements at Castlecreek Plaza were designated as shown in Figure 4. The results of the license plate matches at Castlecreek Plaza are shown in Table 6.

TABLE 4 Results of License Plate Matches at Eastway Plaza Based on Two Trip Types

MOVEMENT TYPE	# OF MATCHES	TRIP TYPE	MOVEMENT TYPE	# OF MATCHES	TRIP TYPE
G→H	17	PB/NPB	B→C	9	PB/NPB
Q→R	3	PB/NPB	B→E	10	PB/NPB
N→O	0	PB/NPB	N→M	1	PB/NPB
L→M	1	PB/NPB	L→O	1	PB/NPB
D→E	1	PB/NPB	D→C	0	PB/NPB
G→R	17	NPB	Q→H	3	NPB
B→M	4	NPB	B→O	7	NPB
L→C	5	NPB	L→E	6	NPB
D→O	0	NPB	D→M	0	NPB
N→E	0	NPB	N→C	0	NPB
G→E	5	PB	G→O	1	NPB
G→C	10	PB	G→M	12	PB/NPB
Q→E	2	PB/NPB	Q→O	0	NPB
Q→C	5	PB/NPB	Q→M	2	NPB
N→R	0	PB/NPB	N→H	0	NPB
D→R	0	NPB	D→H	0	NPB
B→R	5	PB/NPB	B→H	6	NPB
L→R	1	PB	L→H	1	PB/NPB

TABLE 5 Results of Interview Survey at Marsh Supermarket (Conducted by A&F Engineering, Inc.)

TIME	PASS-BY TRIPS	NON-PASS-BY TRIPS
5:00-5:15	4	2
5:15-5:30	5	5
5:30-5:45	6	9
5:45-6:00	6	4
6:00-6:15	5	7
6:15-6:30	6	9
TOTAL (PERCENTAGE)	32 (47%)	36 (53%)

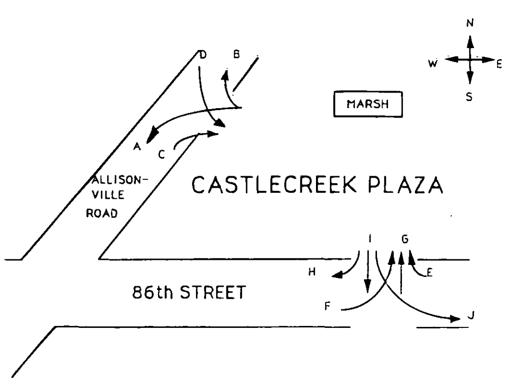


FIGURE 4 Turning movements at Castlecreek Plaza.

As with the case of Eastway Plaza, the researchers used the turning movements at the driveways along with their knowledge about nearby land uses and travel patterns in the vicinity of the site to classify each pair of in/out movements as pass-by, non-pass-by, or both.

For example, movement F most likely came from I-465 approximately 0.5 mi south of 86th Street. If there were easy access for a return to I-465 to the east and south via J, the analysts would have considered FJ either a pass-by or diverted trip, and therefore "PB/NPB" in Table 6. But because trips returning to I-465 usually use Allisonville Road, movement FH is considered a diverted or primary trip (NPB in Table 6), and FJ is listed as pass-by only.

Using knowledge of the surrounding street network and land uses is better than adopting a rigid rule such as "all left-in–left-outs are pass-by trips." For example, movement QR in Figure 2 could be a trip diverted from other north-south routes east or west of US 52.

TABLE 6 Results of License Plate Survey at Marsh Supermarket, Castlecreek Plaza

MOVEMENT TYPE	# OF MATCHES	TRIP TYPE	MOVEMENT TYPE	# OF MATCHES	TRIP TYPE
C→A	0	NPB	E→I	6	PB/NPB
C→B	29	PB/NPB	E→J	0	NPB
C→H	2	NPB	F→A	0	NPB
C→I	1	NPB	F→B	6	PB
C→J	2	PB	F→H	3	NPB
D→A	3	PB	F→I	4	PB/NPB
D→B	22	NPB	F→J	3	PB
D→H	3	PB/NPB	G→A	0	PB/NPB
D→I	5	PB/NPB	G→B	11	PB/NPB
D→J	4	PB	G→H	7	PB/NPB
E→A	0	PB/NPB	G→I	1	PB/NPB
E→B	15	PB/NPB	G→J	1	PB/NPB
E→H	6	PB			

RESULTS

The total number of matches and the number of PB, NPB, and PB/NPB trips for the two study sites are shown in Table 7. PB/NPB trips denote the trips for which no decision could be taken (based on their turning movements) as to whether they were pass-by or non-pass-by.

PB_{max} and PB_{min} were calculated under the assumption that all the PB/NPB trips were pass-by and non-pass-by, respectively. PB_{avg} was calculated as $(PB_{max} + PB_{min})/2$. Therefore, PB_{avg} is the number of pass-by trips when there is an equal likelihood that the PB/NPB trips are pass-by or non-pass-by. The percentage of PB_{avg} trips was found to be very close to the percentage of pass-by trips obtained from the interview survey.

The results can also be explained by the following heuristic argument. The actual number of pass-by trips is a random variable and has to lie between PB_{max} and PB_{min} ($PB_{min} \leq PB_{actual} \leq PB_{max}$). It was assumed during the analysis that the NPB trips have two components, diverted (DI) and primary (PR). The PB/NPB trips therefore have two components, PB/DI and PB/PR. In this case, however, there is no occurrence of PB/PR trips, and hence it is a null set. In the absence of any prior information, the probability of the PB/NPB trips' being pass-by or non-pass-by was set as equal (noninformative prior). Under the assumption that the distribution of PB/NPB trips is symmetric, PB_{avg} gives the minimum variance unbiased estimate of the actual number of pass-by trips.

CONCLUSIONS

The percentage of pass-by trips obtained by a license plate survey in both cases was found to be very close to the results obtained by an interview survey. However, further studies have to be conducted before the methodology can be accepted as a standard procedure for estimating pass-by trips. At a minimum, the procedure can place upper and lower bounds on the percentage of pass-by trips for a particular existing development. The range will be small if the analyst has a thorough knowledge about the travel patterns in the area, if most of the driveway turning movements are unambiguous, or both. The range will also be small if the license plate survey includes the adjacent intersections in addition to the driveways.

This method will be especially helpful (in terms of time and personnel requirements) in determining the percentage of pass-by trips for large developments for which a face-to-face interview would be an intensive effort. Even at rather small sites such as Eastway Plaza and Castlecreek Plaza, providing and positioning personnel to conduct interviews can be challenging. Some businesses have so many customers that two or more interviewers are required to administer even a short face-to-face survey. At less active businesses in a shopping center one person could try to cover several adjacent businesses if the spatial distribution of the stores (distance) allowed the interviewer to move quickly to interview arriving customers at the various business entrances.

At Eastway Plaza 13 interviewers and only 3 license plate recorders were employed. The positioning of the interviewers is difficult to plan and may have to be revised during the course of the data collection. Recording of license plates is much easier to plan and, in extreme cases, can even be automated by the use of audio or videotape recorders.

As the size of the shopping center increases, the disparity between the number of personnel needed to conduct an interview survey and the number required to record license plates at driveways increases. And, of course, it is the larger shopping centers that have the greatest need to be studied.

TABLE 7 Summary of Results

SITE	TOTAL MATCHES	PB	NPB	PB/NPB	PB_{max}	PB_{min}	PB_{avg}	INTER-VIEW
EASTWAY	135	19	75	41	60	19	39.5	
(%)		14	56	30	44	14	29	28.5
MARSH	134	24	28	82	106	24	65	
(%)		18	21	61	79	18	49	47

Another important advantage of a workable license plate survey as a substitute for face-to-face interview survey is that it does not interfere with customers en route to or from the businesses or site being studied. This nonintrusive technique is certainly preferred by the customers and the interviewers. It may also be preferable to the property owner when permission to conduct an on-site survey is requested. In extreme cases when a property owner's consent cannot be obtained the license plate study could be conducted from adjacent or nearby property or public right-of-way.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of Saurabh Mukhopadhyay, visiting assistant professor in the Department of Statistics at Duke University, for his help in the application of heuristic arguments. The authors thank the students at Purdue University who participated in the data collection and the anonymous reviewers for their suggestions to improve the paper.

REFERENCES

1. Kittleson, W. K., and T. K. Lawton. "Evaluation of Shopping Center Trip Types." *ITE Journal*, Feb. 1987, pp. 41–46.
2. ITE, *Trip Generation*. Washington, D.C., 1991.
3. Bochner, B., K. Hooper, and A. Rifkin. Traffic Access and Impact Study for Site Development—An ITE Educational Foundation Seminar, ITE, March 1990.
4. Smith, S. A. A Methodology for Consideration of Pass-by Trips in Traffic Impact Analyses. *ITE Journal*, Aug. 1986, pp. 37–39.
5. Dey, S. S. and J. D. Fricker. "Guidelines for Traffic Impact Analysis of Developments Along State Highways, FHWA/IN/JHRP—92/4, Draft *Report, February 1992.*
6. Dey, S. S. *Analytical Methods to Assess The Impacts of New Developments on Highway Segments*, M.S.C.E. thesis. Purdue University, Lafayette, Ind., Dec. 1991.
7. Moussavi, M., and M. Gorman. Refinement of Procedures Used for Estimating Pass-by Trip Percentages. *ITE Journal*, May 1992, pp. 13–16.

Publication of this paper sponsored by Committee on Transportation Planning Applications.