

# Self-Administered Mailback Household Travel Survey

JOHN P. POORMAN and DEBORAH J. STACEY

## ABSTRACT

The development, application, and results of a low-cost self-administered mailback household travel survey used by the Capital District Transportation Committee (the metropolitan planning organization in the Albany-Schenectady-Troy Standard Metropolitan Statistical Area) are described. The objective of the survey was to update 1965 home interview survey data. Criteria employed in selecting the survey technique were that the technique minimize both cost and staff effort and that the technique produce usable personal and household travel information across a broad range of socioeconomic groups and geographic areas. Key features of the survey included the survey instrument design, the use of bus tokens and maps as incentives to increase response rates among lower income households, and the use of an interactive computer program for data entry. The selection of the survey technique and design of the survey instrument are described. Two pretests were employed to verify the effectiveness of the survey instrument and of the use of incentives. After review of pretest results, a full-scale (approximately 5 percent) sample of 12,500 households in the metropolitan area was drawn and the survey was undertaken in March 1983. Results of the pretests and the full-scale survey are discussed, as well as the procedures used to verify accuracy of responses and to enter data into computer files. Costs of each aspect of the survey effort are also described. The survey successfully elicited detailed socioeconomic and travel data with only a modest level of staff involvement and survey cost. Cost per usable survey response was approximately \$14.00 for 2,610 responses.

In the spring of 1965 a comprehensive origin-destination survey was conducted in the metropolitan core of the Capital District (Albany-Schenectady-Troy, N.Y.) by the New York State (NYS) Department of Public Works (1). The study consisted of three parts: home interview, truck and taxi survey, and external survey. A vehicle-miles-of-travel survey was also conducted that attempted to measure the magnitude of vehicular travel in the survey area.

The Department of Public Works conducted the home interview portion of the survey to determine household, personal, and trip-making characteristics of residents and out-of-town visitors in the Capital District. The survey was taken in the urban (cordon) area and data were collected from one out of every fifty households (or 2 percent) in this area. A total of 3,697 samples were originally selected; however, 3,048 surveys were completed. The truck and taxi survey consisted of selecting a sample of 5 percent of commercial vehicles registered with the NYS Department of Motor Vehicles. The external survey was conducted by setting up roadside interview stations and portable traffic counters at the major and minor highway crossings of the cordon line.

The resulting origin-destination information and trip generation data, by mode and purpose, for various groups of urban households served as the data base for transportation planning in the Capital District through the 1970s. After the 1980 census, the Capital District Transportation Committee (CDTC), as metropolitan planning organization (MPO) for the area, programmed staff effort to investigate the feasibility of a small-scale household travel survey to supplement journey-to-work and demographic information available from the census.

Objectives of the CDTC survey were

1. To determine trip rates by mode, purpose, and occupancy for various classes of households grouped by automobile availability and employment status;
2. To determine median trip lengths by geographic area;
3. To allow for weighting of data according to 100 percent sample census results to provide regional means and medians for trip characteristics; and
4. To allow for detailed analysis of origins and destinations of all personal travel.

The resulting information would serve, with census information, as a data base for continuing transportation planning. Several survey techniques were investigated by the CDTC central staff to determine the anticipated cost and effectiveness of each.

## SURVEY SIZE

Preliminary calculations indicated that 2,500 usable responses from a random sample of households would maintain adequate accuracy at all levels of desired disaggregation.

The 2,500 responses were expected to provide accuracy within  $\pm 7.9$  percent at 95 percent confidence for regional means such as household trip rate (2). The 2,500 responses also would allow for acceptable levels of accuracy for rates calculated for subsets of the regional household population. Using up to nine subregional areas and four socioeconomic classifications to create 36 household "cells," for example, would still provide accuracy for means of household trip rates within approximately  $\pm 25$  percent at the 95 percent confidence level (2). Calculations performed at the trip level (median trip length, for example), rather than at the household level, would have an even greater level of accuracy because they are based on a sample of more than 20,000 trips (at 8+ trips per household).

Adequate response by socioeconomic or geographic cell was deemed important because it was expected that responses would be adjusted to the true distribution (from the 1980 census) of households by cell to produce accurate regional means. A total of 2,500 responses were considered adequate to assure adequate response by cell.

## SELECTION OF SURVEY TECHNIQUE

Several alternative techniques were examined, including traditional home interview of a sample of households, mailback survey distributed to nearly all households by means of an insert in daily newspapers, telephone interview of a sample of house-



holds, and mailback survey mailed to a sample of households. Examination focused on cost and effectiveness comparisons of the techniques as they would be applied to a detailed household and personal travel inventory. Minimizing staff administrative burden was also a consideration because it was desired to undertake the survey largely with existing staff during normal working hours.

### Mailback

The major advantages of the mailback technique for the CDTC survey were its low cost and low administrative burden. It has been estimated that the cost of mail surveys is traditionally 20 percent of the cost of home interview surveys, per completed interview, and approximately 60 percent of the cost of telephone surveys, per completed interview (3).

The low administrative burden was reflected in the fact that the effort could be accomplished largely by existing staff during normal working hours. The survey could be expected to be largely self-administered, with availability to assist respondents by telephone during normal working hours the major obligation of staff members during the survey period. Postcard reminders would be used to increase response rates; evening work would therefore be limited to those callbacks necessary to complete or correct returned surveys for households that indicated that they were unavailable during the day. All other work could be accomplished as staff schedules permitted during normal working hours. As a result of this review of available techniques the Capital District Transportation Committee selected the mailback technique for testing in January 1982.

Some "rules of thumb" with respect to the mailback survey instrument design were followed to ensure usable results. These rules were

- Design the survey instrument so that the question answering sequence has no effect on the accuracy of the elicited responses;
- Design the survey instrument so that the questions asking for objective information can easily be checked or filled in;
- Ask a sponsor who is in good standing with and well known to the target population to sign the introductory letter or postcard and cover letter; and
- Stress the importance of the survey and the confidentiality of the responses in the cover letter, and if possible draw a comparison between the Census of Population and Housing and the travel survey.

A successful Kentucky Department of Transportation (KDOT) survey served as a starting point for the design of the instrument (3). The previously mentioned rules of thumb were used to adapt the KDOT instrument to the broader Capital District objectives of trip information on all trips, rather than information on just automobile driver trips as sought by KDOT. It was also determined that a pretest would be used to verify the feasibility of the technique for the Capital District survey before a commitment to undertake a full-scale survey would be made.

### SURVEY DESIGN AND PRETESTS

#### Survey Design

The survey instrument was designed during the spring of 1982. The instrument was composed of an intro-

ductory letter signed by the Chairman of the Capital District Transportation Committee (a well-known local elected official), a household questionnaire, five travel tables, and a pre-addressed business reply envelope.

The layout was designed to elicit answers to straightforward questions about household size and composition, employment status, vehicle availability, and household income, before tedious travel tables were presented. The income question was asked as the final question in the household questionnaire sequence to maximize respondent cooperation in the survey.

Physically, the household questionnaire formed the outside of a folder containing the five loose, two-sided travel tables. The household questionnaire was designed to detach along perforations so that it could accompany completed travel tables in the return envelope. Instructions for the travel tables were printed on the inside of the front of the folder (the reverse side of the household questionnaire), and reminders and final instructions were printed on the inside of the back of the folder. Instructions were designed to be as brief and as simple as possible, to avoid discouraging potential respondents from completing the forms. Also, instructions included the name and telephone number of a staff member who could be contacted for assistance. Travel tables provided space for information on up to 11 trips per person. Respondents were instructed to use a separate sheet of paper for any additional trips.

Respondents were asked to assign a "person number" to each household member over the age of five and to provide sex, age, relationship to head of household, and employment status data for each member. Respondents were also asked to note whether each member traveled on the designated travel day. During the pretests, a specific midweek day was designated for the travel inventory; however, this required coordination of the printing and mailing of the surveys to ensure that the survey arrived at the household an appropriate time before the designated day. One result of the pretest was to change the designated day to "Next Tuesday," thus providing greater latitude with regard to timing. The household questionnaire also asked for data concerning vehicle availability, household location, and household income. The questionnaire provided a prominent space for respondents to provide a name, telephone number, and time of day for use by the staff in calling back to verify responses.

The layout of the travel tables was designed to minimize confusion on the part of the respondent. Arrows directed the respondent from question to question for each trip. "Checkoff" answers were used to the extent possible to minimize effort.

In addition, the form used a "Then I went to" method for linking trips in such a way as to eliminate the need to duplicate the destination of one trip as the origin of the next trip. The destination of the trip alone was requested. This also eliminated the extra work and confusion caused by asking "Purpose from" and "Purpose to" as was done in the KDOT survey and others.

One travel table was to be completed for each household member 5 years of age or older who traveled during the designated travel day. Respondents were asked to indicate the person number of the household member to whom the travel table applied.

To improve response rates, a follow-up postcard was sent to each household, timed to arrive within 1 week of receipt of the survey. The postcard requested that the householder complete and return the survey forms if he or she had not already done so.



### First Pretest

On June 8, 1982, the first pretest was mailed to 288 Capital District Area residents randomly drawn from telephone directories. Of the 288 questionnaires that were mailed, 30 could not be delivered as addressed. No attempt was made to replace these 30 sample respondents for the pretest. By Friday, June 18, 31 completed surveys had been received, and on June 21, reminder postcards were mailed to the remaining 227 nonresponding households. Of the 258 surveys that were delivered, a total of 54, or 20.5 percent, were completed and returned (31 or 12.0 percent prepostcard, 22 or 8.5 percent postpostcard), and 43 (16.7 percent) of these were usable as returned or after brief verification telephone calls. Eleven unusable responses were from households that could not be reached during the 8 a.m. to 5 p.m. work day and were therefore not contacted.

The results of the first pretest verified that the survey instrument was effective in producing usable results. Mean trip rates, mode share, household size, and household income were very close to expected values. The overall usable response rate of 16.7 percent was nearly at the expected 20 percent level, even without repeated callbacks. No firm conclusions could be reached concerning the response rate by income group because of the size of the pretest sample, but although underrepresentation among lower income groups had been expected, responses from the \$0 to \$10,000 household income class represented 17.1 percent of all usable responses, compared with an expected value of 19.7 percent based on household distribution. The only significant deficiency in response rate was in the \$0 to \$4,999 group; this deficiency could be explained by the telephone directory sampling procedure.

Underrepresentation was found in response rates from two geographic areas: the city of Albany and the suburban area to the southwest of Albany (Delmar and the eastern parts of Rotterdam and Guilderland). Usable response rates from these two areas were 3.7 percent and 3.4 percent, respectively, compared with 16.7 percent across all nine subregions. Although the variation among subregions could not be statistically verified, it was concluded that, if the overall response rate and the variances of the response rates between subregions could be corrected, a full-scale final survey would be warranted. Because most travel data reported by the responding households were sound, some form of respondent incentive, and to a lesser extent the survey instrument, was identified for further research.

### Second Pretest

The second pretest was designed primarily to test the effect of incentives on response rate in areas from which response had been low in the first pretest. The introductory cover letter included an offer of two bus tokens for use on Capital District Transportation Authority (CDTA) buses or a free county map as a token of appreciation for completing the questionnaire. Reference to the incentive was included on the reminder section of the questionnaire folder, and boxes to check for token, map, or "nothing, thank you" were added to the return envelope.

Minor revisions to the survey forms were made to clarify the "How did you get there" column by rephrasing the mode choice options. (For example, "car/van/truck-driver" was changed to "driver-car/van/truck" to properly separate drivers from passengers.)

A clarification of the difference between CDTC and CDTA was made through an introductory postcard. The postcard was intended to improve the response rate among households that were not transit users; several respondents to the first pretest refused to cooperate because "they didn't have any bus service anyway."

During September 1982 a second household survey pretest incorporating the recommendations that resulted from the initial pretest (introductory postcard, offer of incentives, alteration of the "How did you get there" column on the travel table, and inclusion of a space to indicate the completion date of the survey) was conducted. One hundred households were chosen from the two subregions (superdistricts) that had low response rates in the first pretest. Of the 50 surveys mailed to households within each superdistrict, 47 were actually delivered in superdistrict 1 (Albany) and 49 were actually delivered in superdistrict 5 (Delmar and so on).

A total of 22 responses were received, 20 of which were useful. Ten responses were received from households located in superdistrict 1 (9 useful) and 12 responses were received from households located in superdistrict 5 (11 useful). Hence, the useful response rate for superdistricts 1 and 5, respectively, was 19.1 percent and 22.5 percent, which represents a total response rate of 20.8 percent for both districts. Of the 20 responding households, all but 4 requested a token of appreciation: nine respondents requested county maps and seven requested CDTA bus tokens.

The usable response rate for these two superdistricts was judged significantly higher than that obtained in the first pretest (20.8 percent versus 3.5 percent). It was concluded that the usable response rate from a full-scale survey would be at least 20 percent across all superdistricts and that a mailout of the survey to 12,500 households (4.7 percent of the four-county Capital District total of 267,000 households) would produce the desired 2,500 responses. The commitment to undertake the full-scale survey in the spring of 1983 was made by the CDTC Planning Committee in October 1982.

### SAMPLING METHODOLOGY

The success of any sampling methodology depends on the completeness and accuracy of the sampling frame. Area telephone directories were used as the sampling frame for the two pretests but were found to be inadequate because the addresses did not include zip codes, and it was difficult if not impossible to match one of the 100 zip codes in the four-county area to a particular address. Zip code availability was of particular importance for two reasons: (a) accurate zip codes would ensure the best possible delivery and (b) addresses with zip codes would enable the survey and its component pieces (prepostcard and reminder postcard) to be mailed at third-class bulk rates for an estimated savings of \$3,000 over the first-class mail rate. A search for more nearly perfect sampling frames resulted in the decision to use cross-reference directories as the major portion of the sampling frame. Three cross-reference directories (CRDs), the Hill-Donnelly Cross-Reference Directory covering Schenectady and vicinity, the City Publishing Company Cross-Reference Directory covering Greater Troy and vicinity, and the City Publishing Company Cross-Reference Directory covering Greater Albany and vicinity, were purchased (total cost \$190). These three directories, published in April 1982, contain telephone directory listings rearranged by location (street/area) and by telephone number sequence. The listing of households by location was used as the sampling



frame because zip codes were included in this listing. However, four inadequacies of this sampling frame were acknowledged:

- The CRDs are essentially rearrangements of telephone directories; therefore only those persons having listed telephone numbers are included.
- Addresses appear just as they do in the telephone directories, which means that some addresses in the CRDs do not contain specific box or street numbers.
- The CRDs cover only 92 percent of the sampling area, and the sample from the remaining portion of the sampling area had to be chosen from a different sampling frame.
- The CRD listings could not be easily separated into the nine superdistricts, and thus the sample could not be chosen by superdistrict.

These deficiencies in the sampling frame were analyzed before the CRDs were purchased and it was concluded that none would interfere with the choice of the sampling methodology. It was also determined that the representativeness of the sample chosen would not be undermined; samples for many other successful household travel surveys relied solely on telephone directories (4).

Because the sampling frame could not be divided into superdistricts, in accordance with the methodology suggested by David Hartgen, NYSDOT (5), it was decided that a systematic sample would best suit the available sampling frame. It was assumed that this sampling technique would ensure that all areas would be sampled and that areas more heavily populated with households would be more heavily sampled and result in a proportionately accurate sample (one that would reflect the true distribution of households by town, municipality, county, or when aggregated, by superdistrict). For example, if the city of Albany had twice as many households as the city of Schenectady, it was assumed that, by taking a proportionate systematic sample, the sample would contain twice as many Albany city households as Schenectady city households. It was assumed that, after data collection, comparison of results by superdistrict could be made as long as the population of each superdistrict was large enough to have produced a sufficient number of responses.

To ready the CRDs for this type of sampling, the overlap between books was partitioned off, and sections containing addresses of residences outside the sampling area were also blocked off. Similarly, pages containing only businesses were marked. The remaining pages were numbered 1 through 700. It was estimated that the three directories included 91.8 percent of the sampling area household listings and that the required sample size from the directories was 11,475 (12,500 x 0.918). It was calculated that if 17 samples were chosen per page, the total number of samples chosen from the CRDs would be close to 11,475. (Some pages that were not totally excluded contained many business listings. Therefore, to make the sampling fair, that is, so that one household would not have a higher probability of being chosen than another, only 5 to 10 samples were chosen from these pages.)

Based on the average number of household listings per page, it was determined that the choice of every 30th household (businesses were skipped) per page, after starting randomly, would require the sampler to go through each page at least once in order to end up with 17 choices.

Eleven towns in northern Saratoga County were not covered by the CRDs. Household population figures for each of these 11 areas were estimated and their percentage distribution calculated. To be consistent with a proportionate sampling methodology, these percentages were multiplied by the target number of samples, 1,025 (12,500 - 11,475), to estimate the number of samples needed from each town. Area telephone directories were then used to choose the 1,025 samples. Again, a systematic sampling technique was used. One additional step was required, however: each sample address was located on a map that showed zip codes, and each address was assigned a zip code. A total of 12,482 samples were chosen using the methodologies described; 11,447 household names and addresses were chosen from the CRDs and 1,035 from the northern Saratoga County telephone directories.

#### SURVEY PREPARATION FOR MAILOUT

As mentioned previously, low cost was a primary consideration in planning and developing the survey mechanism and associated tasks. To keep costs at a minimum, the persons hired to choose the sample were instructed to write each chosen household name and address on a sheet of paper that was overlaid on an address label template. Thirty names and addresses were contained on one sheet and were ready for copying directly onto self-adhesive address labels. The purpose of this was twofold: first, each sample name and address needed to be written only once and second, potentially high typing costs were not incurred. Hence, this process was both cost- and time-efficient. Three sets of address labels were created from each sheet of 30 names and addresses, so that the introductory postcard, the survey package, and the reminder postcard sent to each potential respondent would be addressed identically.

The survey material was sent directly from the printshop to the Workshop, a local organization that trains and employs physically and mentally handicapped persons. This organization charged \$1,700 to stuff the envelopes with the survey package; affix address labels to each of the 12,482 introductory postcards, 12,482 survey packages, and 12,482 reminder postcards; and sort these three separate sets of 12,482 pieces into sacks according to the rules and regulations of bulk mailing--all very labor-intensive tasks. Workshop personnel were extremely efficient in performing these tasks; they finished ahead of schedule, stored the material until it was scheduled to be mailed, foresaw all potential problems with respect to post office requirements, and delivered each mailout to the post office on schedule.

#### SURVEY RESPONSE AND QUALITY CONTROL

The introductory postcard was mailed via bulk mail on March 14 and the survey package was mailed via bulk mail on March 22. The reminder postcard was mailed via bulk mail on March 29. Although precautions were taken to develop an accurate sampling frame, especially with respect to zip codes, a total of 1,825 survey packages were returned because they were undeliverable as addressed. Of these addresses, 1,606 (88 percent) were from the CRDs and 219 were from the telephone directories. The total number of survey packages that were actually delivered was 10,657.

The first incoming call about the survey was received on March 23, and between March 23 and April 22 a total of 89 calls were received. During this 18-day period, 43 calls were received during the first 6 days, 38 during the next 6 days, and 7 dur-

ing the last 6 days. The 89 calls were categorized by reason. The categorizations and distribution of calls by category appear in Table 1. A total of 2,775 completed questionnaires were received between March 28 and June 8, a period of 73 days, 53 of which were work days. Four days elapsed between the survey package mailout and the receipt of the first

TABLE 1 Categorization of Incoming Calls by Purpose

Nature of Call	Number of Calls	Percent of Total Calls
Wrong person? (Are you sure you want me to fill this out?)	16	18.0
Don't use transit	2	2.2
Need help	34	38.2
Complaint (about length of survey, income question, survey in general)	9	10.1
Refusal to complete survey	14	15.7
Other (need more travel tables, addressed to a deceased person)	14	15.7
	89	99.9 <sup>a</sup>

<sup>a</sup>Does not equal 100.0 due to rounding.

completed surveys. Of the surveys returned, 36.8 percent were received during the first 5 work days that returns were received, and 43.8 percent were received during the second week that returns were received. A total of 80.6 percent were therefore received during the first 10 days that returns were received (see Table 2).

As the completed forms were received, a staff member opened each envelope and reviewed the responses for completeness, consistency, accuracy, and usefulness. Each returned survey was then put into one of three boxes—useful, useless, or callback (time specified). For example, some respondents simply mailed back the complete package or a letter stating the reason for their refusal to complete the form. A small number of respondents totally misunderstood the purpose or instructions and sent back totally unusable data. In these cases it was not expected that a callback would produce usable results,

because it would require a detailed explanation of the survey purpose and instructions and would require respondent recall of all trip-making household members. Some survey forms were completed properly (the sequence of trips made sense, travel data for all persons above age 5 was included, and so forth), but one or all personal data questions were not answered or part of the trip information (e.g., minutes, miles) was omitted. These surveys were determined to be useful and it was thought that to call the respondent about the omissions in the information would be an invasion of privacy because the omissions were probably deliberate.

Other surveys that were received were completed properly but either were not 100 percent complete or were not consistent. For example, many completed travel tables were missing a return trip; other questionnaires received indicated that three persons in the household traveled but only two travel sheets were completed. In other responses it appeared that two people traveled together all day, but one person reported making more trips than the other person. Households returning these types of responses were called back so that the accuracy of the information they provided could be verified. A total of 254 households were telephoned, resulting in 250 useful responses (see Tables 3 and 4).

A total of 2,775 surveys were returned, which represents a response rate of 26 percent. Of these, 160 surveys were categorized as useless and 2,608 as useful. Hence, the useful response rate (out of the total number of survey packages that were actually delivered) was 24.5 percent. The useful responses came from approximately 1 percent of all households in the four-county area.

Survey response was considered satisfactory because usable responses exceeded the objective of 2,500. In addition, the effectiveness of the survey design was demonstrated by the low rate of complaint and refusal calls (23 out of 10,657 delivered packages) and calls for help (34). The fact that less than 10 percent of the responses received were not useful was an indication that the survey instrument was understandable by potential respondents.

TABLE 2 Household Travel Survey Response Time Distribution

	Elapsed Work Days <sup>a</sup>	No. Received	Cumulative No. Received	Percentage of Total Received	Cumulative Percentage of Total
Mailout	0				
First week of returns	4	103	103	3.7	3.7
	5 <sup>b</sup>	23	126	0.8	4.5
	6	132	258	4.8	9.3
	7	368	626	13.3	22.9
	8	387	1,013	13.9 <sup>c</sup>	36.8
Second week	9	493	1,506	17.8	54.6
	10	192	1,698	6.9	61.5
	11	206	1,904	7.4	68.9
	12	100	2,004	3.6	72.5
	13	226	2,230	8.1 <sup>c</sup>	80.6
Third week	14	222	2,452	8.0	88.6
	15	73	2,525	2.6	91.2
	16	49	2,574	1.8	93.0
	17	33	2,607	1.2	94.2
	18	22	2,629	0.8 <sup>c</sup>	95.0
Fourth week	23	82	2,711	2.9	97.9
Fifth week	28	37	2,748	1.3	99.2
Sixth week	33	13	2,761	0.4	99.7
Seventh week	38	6	2,767	0.2	99.9
Eighth week	43	3	2,770	0.1	100.0
Ninth week	48	2	2,772	0.1	100.1 <sup>d</sup>
Ten weeks and more	53	3	2,775	0.1	100.2 <sup>d</sup>

<sup>a</sup>Work days elapsed from the day of the survey package mailout.

<sup>b</sup>The reminder postcard was mailed 5 work days (1 week) after the survey package was mailed.

<sup>c</sup>Percentages of total received during the first, second, and third weeks were 36.8, 43.8, and 14.4, respectively.

<sup>d</sup>Does not equal 100.0 due to rounding.

TABLE 3 Timing of Callbacks

No. of Week-days After Mail-out Survey	No. of Households Called
6	23
7	9
8	29
9	16
10	11
11	19
12	44
13	23
14	21
15	35
16	7
17	12
18	4
	253

TABLE 4 Purpose of Callbacks

Purpose	No. of Callbacks	Percentage of Callbacks
Return trip(s) missing	213	83.9
One or more trips omitted	28	11.0
Miscellaneous	13	5.1
	254	100.0

The use of bulk rate mailings was also satisfactory; most respondents appeared to have received material well within the expected 10-day delivery schedule, as evidenced by the return of over 1,000 completed surveys within 8 working days of the mail-out. If any problem was experienced with bulk mailing, it was with the coordination of the presurvey postcard, survey package, and reminder postcard. In some instances these items were delivered out of sequence, and on occasion the postcards were received but the package was never delivered. This prompted a few telephone calls requesting a package. These problems were minor, however, relative to the cost savings achieved by using bulk rate mailouts.

DATA ENTRY

Several alternative procedures for efficient data entry were investigated before the mailout of the full-scale survey. One option available was to have respondents self-code responses on survey forms designed to be readable by optical scanning equipment. This option was ruled out because of the difficulty of designing a survey form that would be both optically scanable and effective in eliciting usable responses.

The second option considered was to have staff code responses on a separate form that could then be optically scanned. However, certain responses such as household and trip destination addresses could not be easily entered on an optically scanable form without making the form unwieldy. Also, household data and trip data would require separate forms. The cost of designing and printing the forms needed and testing available optical scanning equipment was prohibitively high.

The option selected was to develop an interactive computer program that would allow clerical staff to enter survey data directly into a computer file by supplying answers to questions that appear on the computer terminal screen. The program was written in FORTRAN and tested and refined before the survey mailout. The program and data files were entered and maintained through an account at Rensselaer

Polytechnic Institute (RPI) and accessed by remote terminal at CDTC central staff offices.

The procedure involved examination of the returned survey forms, callbacks, and corrections as necessary before data entry. Actual data entry involved logging in to RPI via telephone and entry of survey responses one household at a time. The program solicited data from the terminal operator by means of commands and questions such as "Enter the 4 digit household number" or "How many people live in the household?" The sequence of questions and commands led the operator through the entire survey response in the same order that items appear on the survey forms; responses to the household questionnaire were entered first, then responses to travel tables were entered.

When a negative response to "Did this person make any more trips?" was keyed in, the program moved on to the next person's trips; when all household members' trips were recorded, the program moved on to the next household. At the end of the terminal session, a paper printout of the data file created during the session was obtained and proofed. Corrections were made and the file added to the master file produced to date.

The results of the interactive data entry were quite satisfactory. Detailed household data and trip data including addresses, times of day, trip length, mode, purpose, and occupancy for 2,608 households and 22,308 trips were entered, proofed, and corrected in an average time of 14 minutes per household. A major reason for the efficiency was that no intermediate coding was required. Data were entered directly from the survey forms reducing the time required and minimizing the possibility of generating error in coding.

Costs to execute the program at RPI averaged \$0.50 per household entered; however, the technique would be suitable for application on any microcomputer with at least 90K memory available for program storage. Resulting data files approached 2 megabytes in size and contained literal address records and coded values for all other responses.

A record format code was appended to each record in the data file to allow variable formats to be used. Variable formats were necessary because of the wide variety among households in the number of people per household and trips per person. The record format codes allowed data summary programs to access the varied record directly, using the codes as instruction about what information was contained in each record.

PRELIMINARY RESULTS

Socioeconomic and household person trip rate data from the 2,608 usable surveys were summarized and analyzed with respect to their compatibility with expected results. This process revealed that the reported mean household size (2.673), mean household income (\$28,405), mean vehicle availability (1.68), and average employees per household (1.24) approached their expected values of 2.678, \$25,843, 1.45, and 1.23, respectively (6). In addition, the household trip rate estimate of 8.5 person trips per household per day and the distribution of household trip rates by household size were close to their expected values (the expected average trip rate was between 7.5 and 9.0) (7).

However, further analysis of the data revealed that households of size one, low-income households, and households with zero vehicles available were underrepresented. Hence, the data file was adjusted or weighted by household size and vehicle availability in an attempt to smooth out the differences between



the reported and expected socioeconomic distributions. Although the survey distributions for some socioeconomic variables were not identical to expected values, the trip rate data for each socioeconomic category are still representative of all households in the category because the minimum acceptable number of responses per category still exists. Therefore the data reported by households in each cell can be used in developing trip relationships by socioeconomic class and in producing regional means and medians.

The survey trip data must be further analyzed before final statistics can be made available and conclusions drawn. However, the results from the analyses of the socioeconomic data and household trip rates indicate that the expectations regarding the effectiveness of the survey instrument, sampling design, mailback technique, and accuracy of the survey data have been met.

#### COSTS OF SURVEY DEVELOPMENT AND APPLICATION

Certain inefficiencies were experienced in development of the survey approach that would not occur in repeat applications. Costs for administration and data entry of the full-scale survey were relatively low, however, indicating that the technique is suitable for repeat applications on limited budgets.

##### Survey Development

The inefficiencies experienced in the development of the procedure related to staff caution in making a financial commitment to the full-scale survey. As a result, considerable time was spent researching alternative techniques and circulating draft survey instruments to other agencies for review and comment. In addition to the two pretests discussed in this paper, a "pre-pretest" was performed with neighbors and relatives to gain advice and input to the survey design. Also, research was performed regarding alternative means of producing updates to Capital District travel behavior information without performing a survey. This effort was made to properly weigh the costs and benefits of undertaking a full-scale survey. Inefficiencies were also caused by staff turnover during the development phase. Overall, research, survey design, and all phases of pretests cost approximately \$19,000 over a period of 18 months, including postage, printing, and indirect costs. The major portion of the cost was technical staff charges and indirect costs.

##### Survey Application

The actual administration of the full-scale survey was performed relatively efficiently as a result of an effective survey instrument that limited the number of unusable responses and callbacks. The interactive computer data entry saved considerable data coding time. Review of responses, including assistance to householders with questions and callbacks to households who returned incomplete responses cost only \$3,700. Data entry by clerical staff and subsequent proofing of all entries by technical staff required an average of only 14 minutes per completed response and cost only \$6,900 in staff time. Also, a local handicapped workshop was employed to stuff, attach labels, and sort the surveys for bulk mailing. The cost of this was only \$1,700 or \$0.14 per survey distributed.

These efficiencies resulted in a modest total cost for the full-scale survey. Overall cost was \$36,600 including \$4,500 for printing of survey materials and maps and purchase of bus tokens; \$6,000 in postage for separate mailouts of presurvey postcard, survey package, reminder postcard and incentive, and for return postage for completed and nondeliverable packages; and \$2,000 in computer time-sharing costs. The survey required 660 hours of technical staff time and 770 hours of clerical staff time. The average cost per completed, usable survey was approximately \$14.00.

These costs do not include costs for further refinement and analysis of survey results. The Capital District Transportation Committee has a separate task scheduled for its 1983-1984 work program to undertake a detailed analysis of trip origins and destinations, for example.

Costs for repeat application of the survey could be expected to be somewhat lower; there would be no start-up costs associated with development of the computer coding technique, and CDTC's recent purchase of a large-capacity microcomputer with hard disk would eliminate most all or all of the computer time-sharing costs. The cost per completed survey for a repeat application could be expected to be approximately \$10.00.

Costs of application of the survey procedure in other metropolitan areas would vary, depending primarily on agency indirect rates. For the CDTC survey, indirect rates of 120 to 160 percent were included in costs of technical staff time, and no indirect charges were associated with temporary, clerical staff time. Table 5 gives the costs associated with the survey application.

TABLE 5 Cost of Survey Application

Task	Person Hours		Cost (\$)			
	Technical	Clerical	Staff	Printing	Postage	Other
Survey formulation and sample design	40	—	1,073	—	—	190 <sup>a</sup>
Sampling	116	300	4,305	—	—	—
Mailout preparation	70	11	2,435	2,300	—	—
Mailout	—	—	—	—	5,300	1,700 <sup>b</sup>
Research coding techniques, develop coding programs	148	—	3,707	—	—	—
Receive incoming calls, review responses, call back respondents	125	90	3,737	—	—	—
Data entry and proofing	162	443	6,914	—	—	2,000 <sup>c</sup>
Incentive mailout	—	—	—	2,250	750	—
Total	661	844	22,175	4,550	6,050	3,980

Note: Dash = no cost.

<sup>a</sup>Directorates purchased.

<sup>b</sup>Staffing envelopes.

<sup>c</sup>Computer time.

## OBSERVATIONS AND CONCLUSIONS

A self-administered household travel survey appears to be satisfactory in eliciting detailed household and personal travel information. Data collected by the Capital District survey will be used by the CDTC central staff and others for several years in updating 1965 travel relationships currently used in travel forecasting and other activities. Combinations of data obtained will also permit analysis of trip-making characteristics of various "life-cycle" groupings of households, and median trip length information by geographic area will be useful in updating traffic simulation models. The technique appears applicable to other metropolitan areas and repeatable in the Capital District for a modest investment of staff and financial resources.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of the New York State Department of Transportation in the development of the survey design and the effort of members of the CDTC central staff, particularly Glenn Posca and June Maniscalco, who designed and carried out the computer data entry procedure. Funding for this planning activity was provided by the Federal Highway Administration and the Urban Mass Transportation Administration as part of the Capital District Transportation Committee's Unified Planning Work Program.

## REFERENCES

1. R.W. Memmott III. Home Interview Survey Manual. New York State Department of Public Works, Albany, N.Y., 1966
2. G.H. Tanner. Allocation of Sample Households for the 1970 Rochester Home Interview Resurvey. New York State Department of Transportation, Albany, N.Y., 1970.
3. Travel Survey Procedures for Statewide Transportation Planning. U.S. Department of Transportation, 1976.
4. Urban Origin-Destination Surveys. U.S. Department of Transportation, 1972.
5. D.T. Hartgen and J.A. Knoll. Design for the Genesee Transportation Study Travel Survey. New York State Department of Transportation, Albany, N.Y., 1974.
6. Characteristics of People and Housing: Summary Statistics. In 1980 Census of Population, New York State Data Center, Albany, N.Y., 1980.
7. G. Cohen and M. Kocis. Components of Change in Urban Travel. In Transportation Research Record 755, TRB, National Research Council, Washington, D.C., 1980, pp. 42-47.

Views of the authors expressed in this paper do not necessarily reflect policies of the Capital District Transportation Committee or its funding agencies.

# Application of the Highway Condition Projection Model to Interstate 4-R Repair

DAVID T. HARTGEN

## ABSTRACT

Procedures developed by the New York State Department of Transportation to evaluate repair strategies for the Interstate Resurfacing, Restoring, Rehabilitating and Reconstructing (I-4R) Program are described. Two procedures were used: (a) 5-year work programs for projected I-4R expenditures, developed by the department's 11 regional offices and based on a preliminary allocation of funds to substate areas, and (b) regional-level quantification of current and projected pavement repair needs using the highway condition projection model (HCPM). Both methods produced generally similar results. The HCPM was generally able to identify sections in need of repair and the work required. Overall, the HCPM placed pavement needs estimates at \$164 million for 5 years; if implemented, these

actions would substantially improve the condition of older New York State Interstates. Regional cost estimates for work needed were higher than HCPM estimates because of included nonpavement improvements. The analysis concludes that an overall network view of repairs is useful in balancing more specific project assessments, which are best prepared by experts closest both administratively and geographically to the project.

Numerous studies have documented the existence of significant deterioration in the extensive system of U.S. roads. At the national level estimates of the repair bill for highways and bridges run upward from \$100 billion (1). Although evidence from the most recent Highway Performance Monitoring Study (2) suggests that the condition of local and state roads is worse than that of Interstates, most recent attention has focused on the overall condition and carry-