TRANSPORTATION RESEARCH

RECORD

No. 1412

Planning and Administration

Innovations in Travel Survey Methods

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1993

Transportation Research Record 1412
ISSN 0361-1981
ISBN 0-309-05559-8
Price: $24.00

Subscriber Category
1A planning and administration

TRB Publications Staff
Director of Reports and Editorial Services: Nancy A. Ackerman
Associate Editor/Supervisor: Luanne Crayton
Associate Editors: Naomi Kassabian, Alison G. Tobias
Assistant Editors: Susan E. G. Brown, Norman Solomon
Production Coordinator: Sharada Gilkey
Graphics Coordinator: Terri Wayne
Office Manager: Phyllis D. Barber
Senior Production Assistant: Betty L. Hawkins

Printed in the United States of America

Sponsorship of Transportation Research Record 1412

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Chairman: Sally Hill Cooper, Virginia Department of Transportation

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
David Preston Albright, Marsha Dale Anderson, Jack A. Butler, Patrick R. Cain, Ed J. Christopher, Keith A. J. Crawford, Joyce Curtis, 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, George F. Wiggers

Committee on Traveler Behavior and Values
Chairman: Ryuichi Kitamura, University of California–Davis
Elizabeth Deakin, Bernard Gerardin, Thomas F. Golob, David T. Hartgen, Joel L. Horowitz, Lidia P. Kostyniuk, Martin E. H. Lee-Gosselin, Hani S. Mahmassani, Patricia L. Mokhtarian, Frank Southworth, Peter R. Stopher, Janusz Supernak, Antti Talvitie, Mary Lynn Tischer, Carina Van Knippenberg, Martin Wachs, William Young

James A. Scott, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1992.

Transportation Research Record 1412

Contents

Foreword

The papers in this volume present a variety of innovative techniques to monitor, collect, and disseminate travel survey information. Several papers focus on techniques of collecting travel information, such as travel diaries, origin-destination surveys of households and expansion weighting procedures to generate population estimates, factoring procedures to eliminate nonresponse bias, a two-staged household survey mail-back questionnaire, and an analysis of underreporting of trips in telephone interviews.

Also in this Record are papers reporting the results of video and computer applications. Video imaging technology to collect flow rate data is described; the effective use of video when roadside interviews are not possible is discussed; and the practice of image processing for the segmentation and matching of vehicles in road images is explained. Computer applications include a survey system that uses a touch-screen interface to elicit data on user satisfaction and a data acquisition system that employs tape switch sensors on the roadway. The overall effectiveness of entering information from a telephone interview survey directly into a computer file is also studied.

Study of Car Travel Characteristics in Singapore

T. F. FWA, B. W. ANG, AND T. T. NG

A survey conducted in 1990 to study the car travel characteristics in Singapore, a city-state with a population of about 3 million, is described. A vehicle-based travel-diary approach was adopted in which the details of all trips made by a sample of 115 cars over an average period of about 4 weeks were collected and analyzed. A total of 11,638 trips were recorded in the survey. Trips were cross-classified and examined by trip purpose, time of travel, and road type. Travel characteristics such as the temporal pattern of car travel, trip duration, trip distance, occupancy level, and trip speed were analyzed, and travel patterns of cars on weekdays and on Saturdays and Sundays were examined. A least-squares fitting technique was applied to provide estimates of journey speeds on different road functional classes. The approach adopted has the advantage that the data offer accurate records of travel by participating cars in actual traffic conditions. The results of the study indicate that the approach could be used as a useful alternative to provide information on car travel characteristics.

Information on travel characteristics of car owners is vital to the planning of transportation systems and formulation of traffic management schemes. Transportation engineers and planners can use such information to analyze existing traffic flow conditions and to evaluate the possible effects of transportation-related policies. Home interview surveys have been a commonly used technique for obtaining relevant data on trip-making characteristics of households, car owners, or adults. The transportation agencies of Singapore have conducted several large-scale household travel surveys in the past 12 years (*1–4*). These household surveys provided useful information on car ownership, distribution of person trips by trip purpose and by mode of travel, and spatial variations in trip patterns.

Certain important travel information—such as vehicle speed by road class and by time of day, trip distribution by road class, and length of trip by road class—is not available from these surveys. Travel speed data are particularly valuable as an indicator parameter of the effectiveness of various traffic management schemes that have been implemented in the island city-state of Singapore. In view of this, a car-travel diary survey was undertaken in 1990 to study the car travel characteristics in Singapore that could not be derived from the earlier household surveys. This paper presents the methodology and results of this study.

Center for Transportation Research, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511.

DESIGN OF SURVEY

Basis of Survey Design

A vehicle-based approach was adopted in which the details of all trips made by a sample of 115 cars over an average period of about 4 weeks were collected and analyzed. A major objective of the study was to identify the travel characteristics of survey participants in actual traffic conditions over the study period. Another objective of the overall project, which has been discussed elsewhere (*5*), was to develop automobile fuel consumption models that related fuel consumption to trip length, vehicle speed, and vehicle occupancy level. The use of fuel-purchase cycles was found to be a convenient basis on which to design the survey to meet the requirements of both objectives.

Survey participants were asked to record vehicle usage on a trip-by-trip basis continuously in three consecutive fuel-purchase cycles. Each fuel-purchase cycle was to cover a minimum total travel distance of 300 km (186 mi). For most participants, a cycle could be completed in 1 to 2 weeks. Each participant therefore took 3 to 6 weeks to complete the survey. Each participant was paid S$100 (approximately $60, U.S. dollars), so the project had a budget to engage 120 participants.

The survey was conducted in two phases: Phase 1 ran from January through April 1990, and Phase 2 ran from August through November 1990. Sixty participants of Phase 1 were selected randomly from people working in the National University of Singapore or whose work required visits to the university to ensure close monitoring by the project investigators (i.e., the authors). Besides receiving written survey instructions, every participant in this phase was given a personal briefing by the project investigators. Phase 2 participants were selected randomly from lists of non-university staff recommended by registered commercial companies and Phase 1 participants. Of 60 people selected to participate, 55 completed the survey.

Survey Details

The survey requires considerable effort on the part of participants to record the details of every trip. Great emphasis was therefore placed in the selection of survey participants to ensure that survey data of good quality were returned. It was on this ground that the target population of survey was restricted to car owners who had at least a high-school education and who drove their cars themselves. The study included a random sample of 115 willing participants. The sample was believed to be representative of the middle-income car owners of Singapore. It is important to note that the participants' day-to-day travel patterns and vehicle use during the period of survey were not affected as a result of participating in the survey.

The following table compares the distribution of engine capacity of the vehicles included in the study with that of the Singapore car population for 1990 (*6*):

Engine Capacity (cm^3)	*Present Study (%)*	*Car Population in Singapore (%)*
Less than 1000	25.2	15.4
1001 to 1600	70.4	65.6
1601 to 2000	4.4	14.7
More than 2001	0.0	4.3

It is seen that cars with engine capacities higher than 1600 cm^3 (98 $in.^3$) were underrepresented and cars with engine capacities smaller than 1000 cm^3 (61 $in.^3$) were overrepresented in the study sample.

Each participant received a book of trip-diary forms and a briefing on how the trip diary should be filled. The trip data to be recorded included the date, the times and odometer readings at the start and the end of each trip, trip purpose, number of adults and children carried, and estimates of the proportions of trip distance by type of road taken. The raw data for each trip were available in the form shown in Figure 1.

Roads were classified into three types: expressways, streets in the central business district (CBD) or satellite residential towns, and other roads not belonging to either of the former types. Specifically, roads classified as "other" refer to major roads and collectors that provide links between satellite towns, between satellite towns and the CBD, and between satellite towns or the CBD and expressways. Three trip purposes were identified: work, business, and miscellaneous. Trips classified as miscellaneous included all trips related neither to work nor to business, such as recreation, social, and shopping trips. Participants were instructed to record the purpose of each using the "to/from" criterion. This simplified representation of trip purpose was adopted as a compromise to reduce the number of data entries in data processing and the number of parameters in the travel speed analysis.

PROCESSING OF SURVEY DATA

A total of 11,638 trips were recorded in the survey. With an average of about 10 data entries per trip, the total number of entries was about 116,000. Considerable time and effort were devoted to data checking and processing. Trip records with missing or illogical data were discarded. The total final number of trip records used in the analysis was 11,589. The relatively low percentage of faulty records was believed to be the result of care taken in survey planning and execution.

Analyses of the data involved cross classifying the number of trips recorded in the survey according to trip purpose, road functional class, and time period of a day. Figure 2 graphically illustrates this method of decomposition of the trip data.

ANALYSIS OF TRAVEL CHARACTERISTICS

The findings of this study pertain to the metropolitan characteristics of Singapore for 1990. Singapore is an island city-state with a land area of 633 km^2. On the basis of the 1990 census (*7*), the population is 2.77 million and the total number of households is approximately 660,000. The average household size is 4.2, and 39 percent of the households are car-owning households. The total car population at the end of 1990 was 272,808, of which 247,808 were household-owned (*6*).

Travel Characteristics by Trip Purpose

Distribution by Trip Purpose

Table 1 presents the proportions of trip by trip purpose. Nearly two-thirds of the trips made by the car owners were for miscellaneous purposes, and about 30 percent were work trips. In terms of person trips, the share of miscellaneous trips increased to 71.6 percent because of higher occupancy levels in such trips.

Table 1 also gives the corresponding proportions of person trips made by car-owning households in Singapore (*4*). Trips by households included trips by public transport such as buses

Date		Start of trip		End of trip	
	Time	______ . ______	am/pm	______ . ______	am/pm
___/___/___	Odometer	__________	km	__________	km

Trip and Road Characteristics

No. of Passengers : Adults ________ Children ________

Road (% by distance) : Expressway ______ CBD/Town ______ Others ______

Trip Purpose : Work ______ Business ______ Miscellaneous ______

FIGURE 1 Type of data that each survey participant was required to provide for every trip.

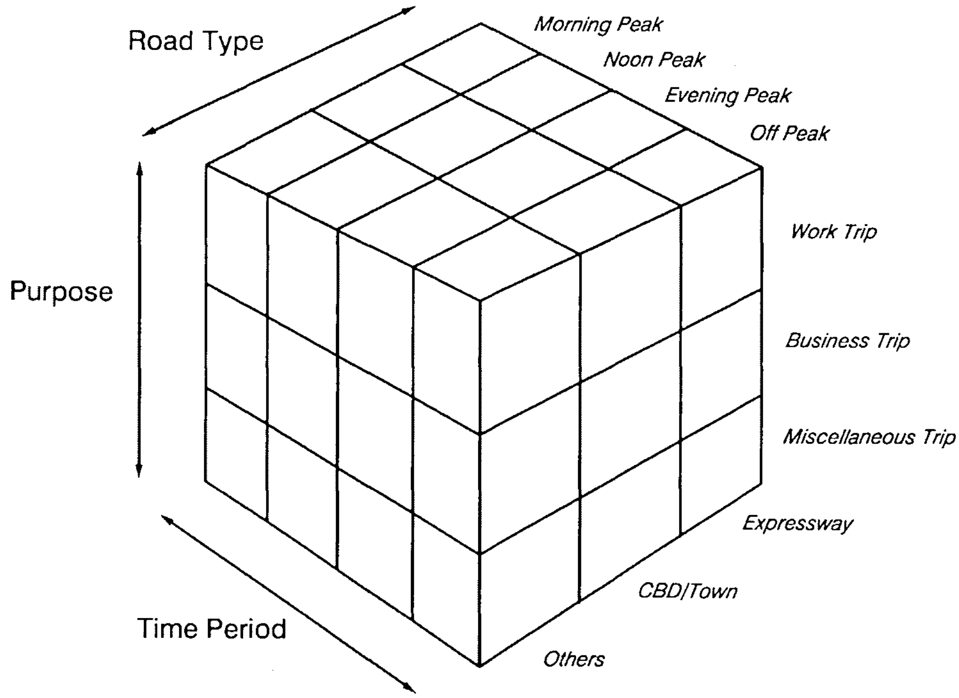

FIGURE 2 Decomposition of trip data.

and rapid mass transit. Comparison of the trip purpose statistics of car owners and car-owning households suggests that a large percentage of the members (including owners of cars) of car-owning households in Singapore did not travel by car to work. The statistics also show that the same was true for business trips. This could be due to several factors unique to Singapore: (a) an area licensing scheme restricts traffic from entering the central city area, (b) the efficient island-wide public transportation system provides a high level of mobility, and (c) Singapore is small in comparison with other major cities, so destinations are easily accessible within reasonable time by most modes of transport.

Time Distribution of Travel

Variations in the times that trips took place were analyzed on the basis of the departure time of each trip. Figure 3 (*top*) plots the distribution of trips by time of day; also shown is the distribution of overall person trips (inclusive of all trips made using different modes of transport). Both distributions exhibit a high morning peak, a slightly lower evening peak, and a much lower noon peak. However, one can also observe the following differences:

1. In the morning, the overall person trips peaked between 6:30 and 8:30 a.m., whereas trips by car peaked later, between 7:30 and 8:30 a.m.
2. For travels by car, the evening after-work peak was followed closely by another peak between 8:00 and 9:00 p.m. This later peak comprised largely miscellaneous trips.
3. The proportion of trips made after 5:00 p.m. was about 46 percent for travel by cars but only about 29 percent for overall person trips.

Most people in Singapore work a half-day on Saturday. This is clearly reflected in Figure 3 (*middle*) by a morning peak between 7:30 and 8:30 a.m. and a higher midday peak between 1:00 and 2:00 p.m. Trips made after 5:00 p.m. still contributed about 40 percent of all trips made by car on Saturday. Figure 3 (*bottom*) shows that there were no pronounced peaking periods on Sunday. Trips were spread evenly between 9:30 a.m. and 9:30 p.m.

On weekdays, the peak-hour car travel in the morning peak period contributed an average of 13.7 percent of total daily

TABLE 1 Distribution of Trips by Trip Purpose

Trip Purpose	Trip Data of Current Study		Person-Trips by Car-Owning Households [4]
	Trips by Car	Person-trips by Car	
Work	28.8%	21.8%	41.5%
Business	8.0%	6.6%	17.9%
Miscellaneous	63.2%	71.6%	40.4%

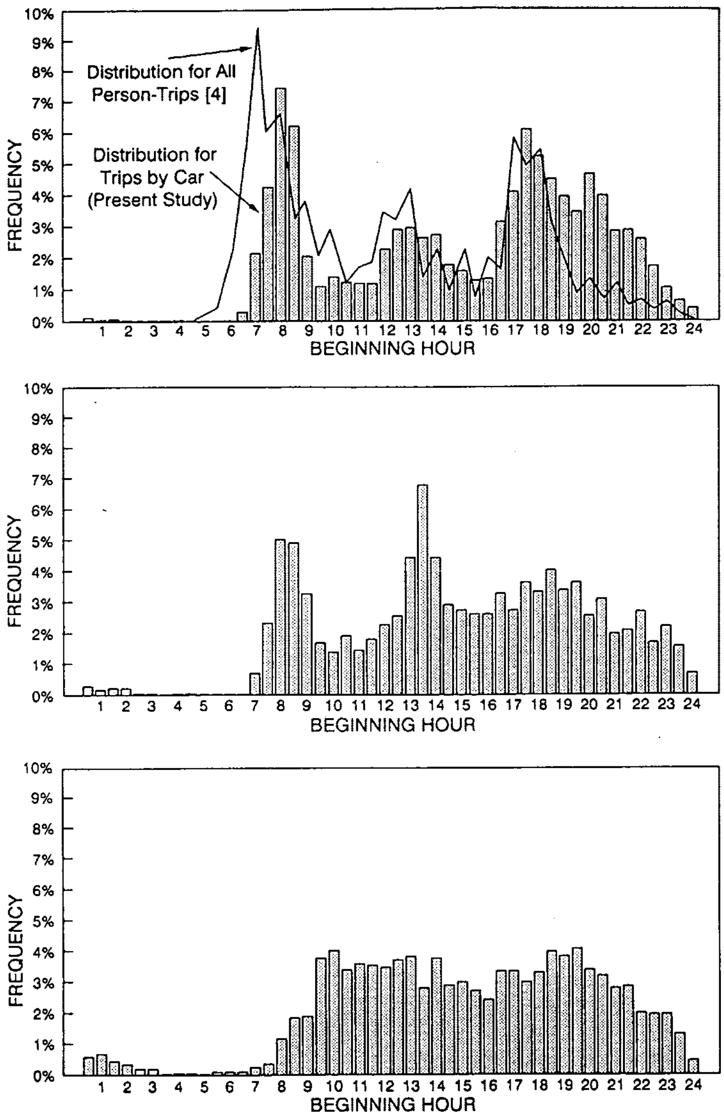

FIGURE 3 Trip distributions by time of day: *top*, time distribution of car travels for weekdays; *middle*, time distribution of car travels for Saturdays; *bottom*, time distribution of car travels for Sundays and public holidays.

car trips. The corresponding percentages for the evening and noon peak periods were 11.5 and 6.1, respectively. On Saturdays, the peak-hour flow in the noon peak period represented 11.2 percent of the daily car trips, and that in the morning peak period had a share of 9.9 percent.

The time distributions of trips by trip purpose are plotted in Figure 4. Figure 4 (*top*) shows that trips to and from work peaked from 7:00 to 9:00 a.m. and 5:30 to 7:30 p.m. Most business trips took place between 11:00 a.m. and 3:00 p.m. It is interesting to see that miscellaneous trips exhibited two peaks: a smaller peak from 12:30 to 2:30 p.m. and a higher and longer-duration peak from 5:30 to 10:30 p.m. The smaller peak coincides with the midday lunch break, whereas the higher peak is probably caused by the after-work social and shopping activities in the evening. Figure 4 (*bottom*) reveals that the morning peak travel was dominated by work trips and the evening peak was caused by both home trips from work and miscellaneous trips. The noon peak was attributed mainly to miscellaneous trips.

Characteristics of Trip Duration

The statistics of trip duration summarized in Table 2 indicate that trips to and from work had the longest mean duration and that miscellaneous trips had the shortest mean duration. This trend is also commonly observed in other major cities in the world (*8,9*). The cumulative curves plotted in Figure 5 demonstrate the same relative magnitude of trip duration by trip purpose. This figure indicates that about 80 percent of car trips were shorter than 30 min. When examined with respect to time of day, Table 2 reveals that trips made during

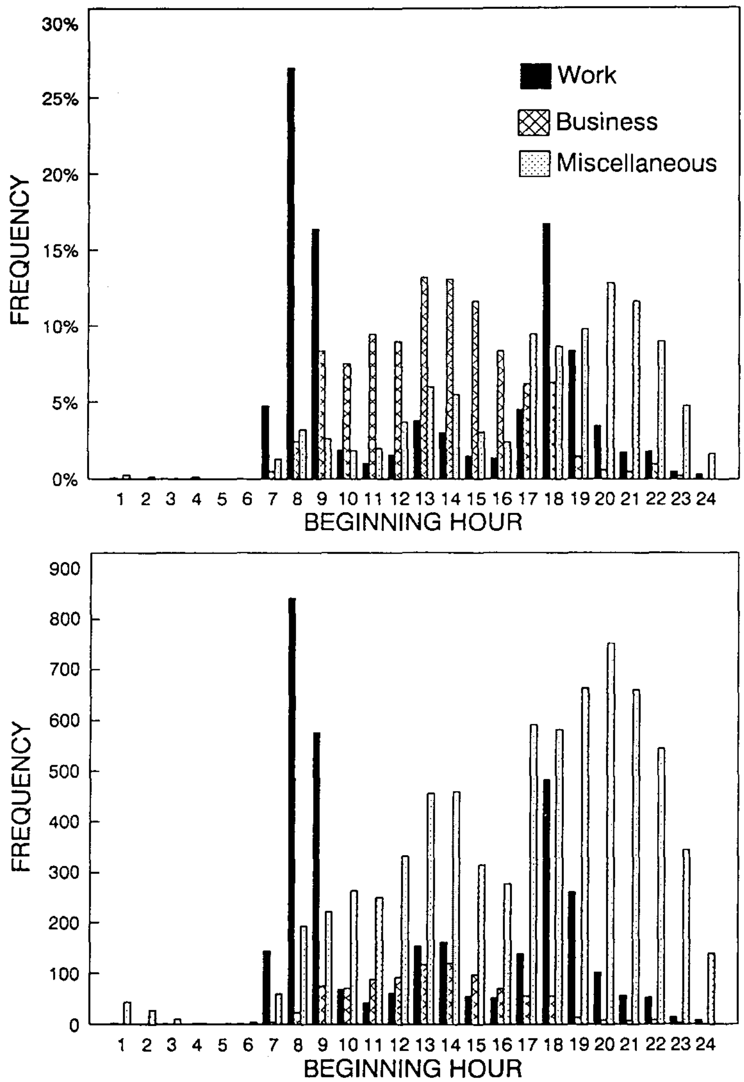

FIGURE 4 **Time distributions for weekdays by trip purpose:** ***top*****, by percentage frequency;** ***bottom*****, by trip frequency.**

TABLE 2 Variations of Trip Duration with Trip Purpose and Time of Day

	All Trips	Trip Purpose			Time of Day(*)			
		Work	Business	Misc.	Peak 1	Peak 2	Peak 3	Off Peak
Mean (min.)	18.7	21.7	18.6	17.2	22.2	14.6	18.8	18.0
Std Dev (min)	13.5	12.8	13.4	12.7	13.6	11.8	12.1	14.1

Note: (*) Peak 1 = morning peak, 7.00 am to 9.00 am,
peak 2 = noon peak, 12.00 pm to 2.00 pm,
peak 3 = evening peak, 5.30 pm to 7.30 pm.

the morning peak had the longest mean duration, those made during the noon peak had the shortest mean duration, and those made during off-peak periods and the evening peak had mean values about the same as the mean duration for all trips.

Characteristics of Trip Distance

It can be seen from Table 3 that the trends of variation of trip duration with respect to trip purpose and time of day, respectively, as described in the preceding section, also hold for trip distance. That is, by trip purpose, the mean distance covered by trips to and from work was the longest and the mean distance covered by miscellaneous trips was the shortest; by time of a day, trips in the morning peak had the longest mean distance and those in the noon peak had the shortest mean distance. Figure 6 depicts the cumulative distributions of trip distance by trip purpose. It shows that about 80 percent of car trips were shorter than 20 km (12.4 mi).

The variation pattern by trip purpose described above was also observed in U.S. cities, with trip distance ratio of work to all trips varying from 1.34 to 1.54 (*8*). In Table 3 the corresponding ratio for car travels in Singapore is only 1.20. Besides, the mean trip distance of 10.9 km (6.8 mi) is also longer than the range of 4.7 to 9.5 km (2.9 to 5.9 mi) reported for U.S. cities (*8*).

Occupancy Level

Table 4 presents the occupancy levels by trip purpose and by time of day. The overall mean occupancy level of 1.75 is higher than the occupancy levels reported for cities in the United States [between 1.4 and 1.6 (*8*)] and Europe (*8*). The higher car utilization is expected because of lower car ownership and higher car usage costs in Singapore (*10*). Work trips had the lowest mean occupancy level and miscellaneous trips had the highest mean occupancy level, a trend also observed in other cities in the world (*8,9*). Off-peak periods were found to register the highest mean occupancy level because of the presence of a relatively high proportion of miscellaneous trips. Trips made during the morning peak (mostly work trips) had the lowest mean occupancy level.

Characteristics of Trip Speeds

The survey data also allow trip speeds to be analyzed by trip purpose and time of day. Table 5 demonstrates that the overall average trip speed was 32.3 km/hr (20.0 mph), which is in good agreement with the recent estimate of 30 km/hr (18.7 mph) made by the Public Works Department of Singapore (*11*). Work trips had the highest mean trip speed, and miscellaneous trips had the lowest mean trip speed. In general,

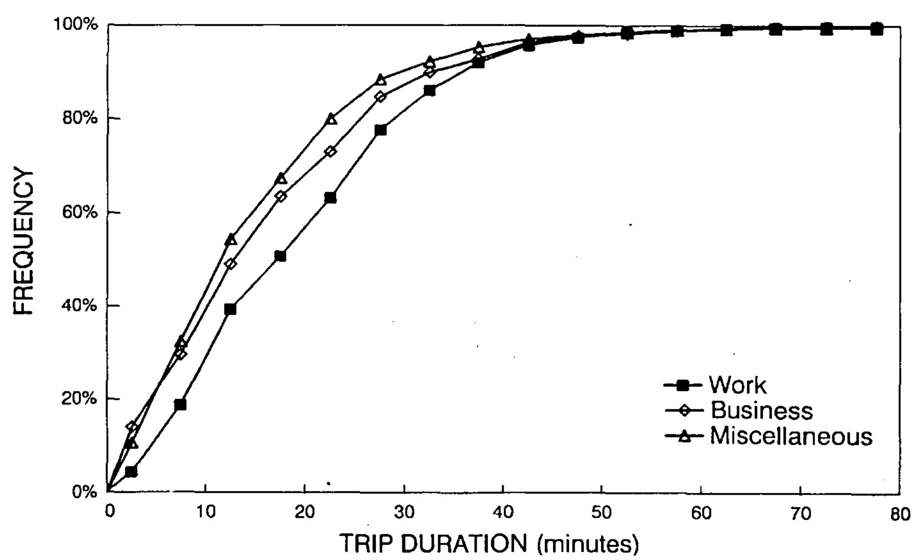

FIGURE 5 Cumulative curves of trip duration by trip purpose.

TABLE 3 Variations of Trip Distance with Trip Purpose and Time of Day

	All Trips	Trip Purpose			Time of Day(*)			
		Work	Business	Misc.	Peak 1	Peak 2	Peak 3	Off Peak
Mean (km)	10.9	13.1	11.1	9.7	13.1	8.3	10.8	10.2
Std Dev (km)	10.5	9.4	9.3	9.6	9.9	8.6	8.9	11.5

Note: (*) Peak 1 = morning peak, 7.00 am to 9.00 am,
peak 2 = noon peak, 12.00 pm to 2.00 pm,
peak 3 = evening peak, 5.30 pm to 7.30 pm.

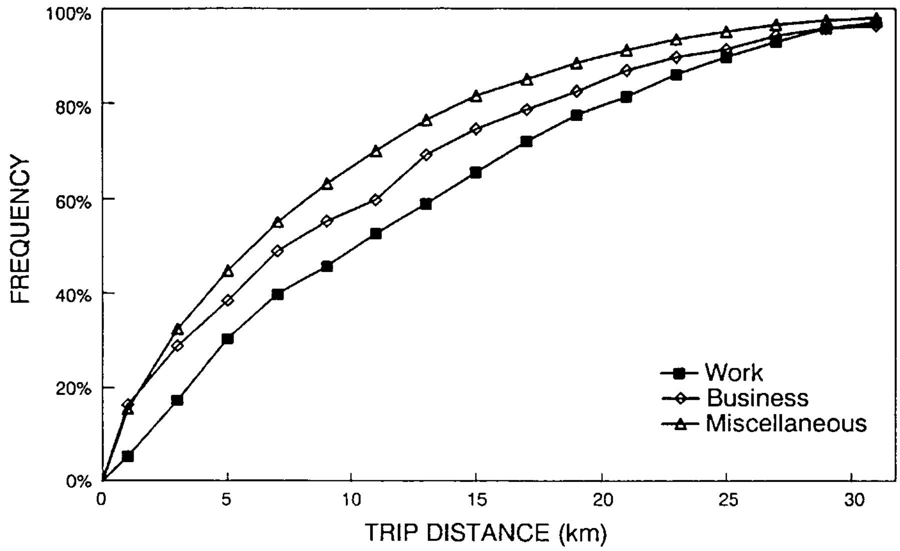

FIGURE 6 Cumulative curves of trip distance by trip purpose.

there was little difference in mean trip speeds among journeys made at different times of day, which could be due to the success of the area licensing scheme that effectively controls the volume of traffic within the CBD during the morning and evening peak periods.

Travel Characteristics by Road Type

Choice of Road Type by Travelers

Most of the trips recorded in the survey covered more than one of the three road types listed in the survey form. Two methods are used to assess the choice of road type by car drivers for their trips: one is to calculate the usage of each road type on the basis of the number of trips that had used it; the other is based on the total distance made on each road type. Table 6 presents the results for both methods of assessment. On the basis of trip usage, about two-thirds of all trips had used the road type "other" in part of the travel, nearly half of all trips covered streets in the CBD or satellite towns, and approximately one-third traveled on expressways. By trip purpose, the same pattern of road use maintained, except that miscellaneous trips differed slightly from the other two trip types by involving more CBD and town streets and fewer expressways.

TABLE 4 Variations of Occupancy Level with Trip Purpose and Time of Day

	All	Trip Purpose			Time of Day(*)			
	Trips	Work	Business	Misc.	Peak 1	Peak 2	Peak 3	Off Peak
Mean	1.75	1.33	1.46	1.98	1.43	1.75	1.52	1.93
Std Dev	1.04	0.60	0.83	1.15	0.69	1.08	0.84	1.15

(*) Peak 1 = morning peak, 7.00 am to 9.00 am,
peak 2 = noon peak, 12.00 pm to 2.00 pm,
peak 3 = evening peak, 5.30 pm to 7.30 pm.

TABLE 5 Variations of Trip Speed with Trip Purpose and Time of Day

	All	Trip Purpose			Time of Day(*)			
	Trips	Work	Business	Misc.	Peak 1	Peak 2	Peak 3	Off Peak
Mean (km/h)	32.3	34.6	33.3	31.1	33.4	30.5	32.4	32.5
Std Dev (km/h)	14.6	13.9	15.5	14.6	14.4	15.3	14.0	14.9

(*) Peak 1 = morning peak, 7.00 am to 9.00 am,
peak 2 = noon peak, 12.00 pm to 2.00 pm,
peak 3 = evening peak, 5.30 pm to 7.30 pm.

TABLE 6 Relative Usage of Road Types by Number of Trips and by Travel Distance

Road Type	Percentage Share in Road Network Mileage	Percentage of Trips That Used the Road Type				Percentage of Total Vehicle-km traveled			
		All Trips	Work Trips	Busin. Trips	Misc. Trips	All Trips	Work Trips	Busin. Trips	Misc. Trips
Express-way	3.8	35.1	40.0	40.9	31.9	40.7	45.8	46.4	36.8
CBD/Town	28.9	46.7	44.6	40.3	48.4	19.6	16.9	19.1	21.2
Others	67.3	66.5	63.6	67.3	67.9	39.7	37.3	34.5	42.0

In terms of distance traveled in vehicle kilometers, expressways and "other" roads had about equal shares of 40 percent each, and CBD and town streets had 19.6 percent. When analyzed by trip purpose, expressways were used most in work and business trips, chalking up about 46 percent in each. On the other hand, the highest percentage of vehicle kilometers traveled (42 percent) for miscellaneous trips took place on "other" roads. Table 6 also indicates that although expressways contributed only 3.8 percent in length to the entire road network of Singapore, they were an important factor in the land transportation system by virtue of the high percentages of traveled distance and trips they attracted.

Derived Characteristics by Road Type

Traveling speed by road type provides useful information on the operating performance of different road types. This information cannot be obtained directly from the survey data because most of the trips covered more than one road type. However, this section shows that the mean journey speed on each road type can be estimated by fitting the survey data to an equation that relates a trip speed to the journey speeds on different road types covered by the trip.

Consider a general case in which a trip covers all three road types considered in the survey. The following information is available from the survey record:

D = total trip distance
d_1 = distance traveled on Road Type 1
d_2 = distance traveled on Road Type 2
d_3 = distance traveled on Road Type 3
T = trip duration

By definition, trip speed V is given by

$$V = D/T \tag{1}$$

Let t_1, t_2, and t_3 be the travel time spent on Road Types 1, 2, and 3, respectively, and

$$\frac{1}{V} = \frac{t_1 + t_2 + t_3}{D} = \frac{t_1}{d_1}\left(\frac{d_1}{D}\right) + \frac{t_2}{d_2}\left(\frac{d_2}{D}\right) + \frac{t_3}{d_3}\left(\frac{d_3}{D}\right) \tag{2}$$

that is,

$$\frac{D}{V} = \frac{1}{v_1}(d_1) + \frac{1}{v_2}(d_2) + \frac{1}{v_3}(d_3) \tag{3}$$

where v_1, v_2, and v_3 are the journey speeds on Road Types 1, 2 and 3, respectively, and the only unknowns in Equation 3.

Since V, D, d_1, d_2, and d_3 are known for every single trip, Equation 3 can be used to estimate the values of v_1, v_2, and v_3 by fitting the 11,589 sets of trip data from the survey records using the method of least squares (*12*). In the current study, the technique of linear regression (*13*) was used to obtain the least-squares estimates for $1/v_1$, $1/v_2$, and $1/v_3$. The dependent variable is D/V, and the independent variables are d_1, d_2, and d_3. The resulting regression expression is

$$D/V = 0.0199(d_1) + 0.0356(d_2) + 0.0234(d_3) \tag{4}$$

The coefficient of determination, R^2, for Equation 4 is .87, and the three coefficient estimates are statistically significant at the 99.9 percent confidence level. The estimates of overall mean journey speed by road type are given by the reciprocals of the coefficient estimates: 50.3 km/hr (31.3 mph) on expressways, 28.1 km/hr (17.5 mph) in the CBD/town, and 42.7 km/hr (26.5 mph) on "other" roads. These estimates are in general agreement with spot speed measurements reported by Singapore road authorities for expressways [ranging from about 40 to 85 km/hr at various times of the day (*14*)] and journey speed measurements for CBD streets [varying between 14.2 and 38.9 km/hr (*15*)].

CONCLUSIONS

This paper describes a study of car travel characteristics in Singapore that was based on travel records of vehicles in actual traffic conditions. The findings obtained from this study are summarized in the following:

- On weekdays, car trips in Singapore peaked from 7:30 to 8:30 a.m., beginning about an hour later than all person trips. A slightly smaller evening peak took place between 5:30 and 6:30 p.m. There was also a small noon peak between 1:00 and 2:00 p.m. The peak hour travel in the three peak periods represented 13.7, 11.5, and 6.1 percent of daily car trips, respectively.
- On weekdays, work trips dominated the trips by car in the morning peak. In the evening peak, there were approximately equal proportions of trips home from work and miscellaneous trips.
- On Saturday, the highest peak occurred between 1:00 and 2:00 p.m. This peak-hour flow provided 11.2 percent of

the daily car trips, compared with 9.9 percent contributed by the peak-hour flow in the morning peak period. No pronounced peaks were observed on Sunday.

- More than 80 percent of car trips were shorter than 30 min. The mean trip duration was 21.7 min for work and 17.2 min for miscellaneous trips. The mean trip duration was longest in the morning peak (22.2 min) and shortest in the noon peak (14.6 min).
- More than 80 percent of car trips were less than 20 km (12.4 mi) in distance traveled. Work trips had the longest mean distance, 13.1 km (8.1 mi), and miscellaneous trips had the shortest mean distance, 9.7 km (6.0 mi). The mean trip distance was longest in the morning peak and shortest in the noon peak.
- The overall mean occupancy level was 1.75. Occupancy level was lowest (1.33) for work trips and highest (1.98) for miscellaneous trips. In terms of time, the morning peak had the lowest occupancy level, 1.43, whereas off-peak periods had the highest level, 1.93.
- Trip speed was highest for work trips and lowest for miscellaneous trips. The mean journey speeds estimated using the least-squares fitting technique were 50.3 km/hr (31.3 mph) for expressways, 28.1 km/hr (17.5 mph) for CBD/town streets, and 42.7 km/hr (26.5 mph) for "other" roads.
- By road type, about 40 percent of vehicle kilometers of car travel were made on expressways and "other" roads; only about 20 percent were made in the CBD/town. In terms of car trips, nearly half involved CBD/town streets, two-thirds involved other roads, and about one-third involved expressways.

REFERENCES

1. *Republic of Singapore Comprehensive Traffic Study—Phase A, Final Report.* Wilbur Smith and Associates Inc., Singapore, 1981.
2. *Report of Traffic Survey in Singapore.* Synergy Business Research and Consulting Pte Ltd., Singapore, 1989.
3. Olszewski, P., and Y. D. Wong. Car Travel and Parking in Singapore—Some Findings from the 1988 Household Travel Survey. *Journal of the Institution of Engineers*, Singapore, Vol. 32, No. 2, 1992.
4. Tan, Y. W., Y. D. Wong, and H. Fan. Trip Characteristics of Households in Singapore. *Proc., 7th Conference of Road Engineering Association of Asia and Australasia*, Vol. 1, Singapore, June 1992.
5. Ang, B. W., T. T. Ng, and T. F. Fwa. A Factorization Analysis of Automobile Fuel Consumption in Actual Traffic. *Energy—The International Journal* (in press).
6. *ROV Annual Report 1990.* Registry of Vehicles, Singapore, 1991.
7. *Census of Population 1990.* Department of Statistics, Census of Population Office, Singapore, 1992.
8. Homburger, W. S., L. E. Keefer, and W. R. McGrath. *Transportation and Traffic Engineering Handbook*, 2nd ed. Prentice-Hall, Englewood Cliffs, N.J., 1982.
9. *Car Ownership and Use.* Road Research Group, Organization for Economic Cooperation and Development, Paris, France, 1982.
10. Oh, S. T., and B. W. Ang. Energy Use and Efficiencies of Singapore Land Transportation. *IES Journal*, Vol. 27, No. 4, Institution of Engineers, Singapore, 1987, pp. 31–38.
11. *The Road Ahead—A PWD Report.* Public Works Department, Singapore, 1992.
12. Sokolnikoff, I. S., and R. M. Redheffer. *Mathematics of Physics and Modern Engineering*, 2nd ed. McGraw-Hill Book Company, New York, 1966.
13. Neter, J., W. Wasserman, and M. H. Kuther. *Applied Linear Statistical Methods*, 2nd ed. Richard D. Irwin, Inc., Homewood, Ill., 1985.
14. Fan, H., P. Olszewski, and D. J. Turner. Analysis of Traffic Flow on Pan Island Expressway. *IES Journal*, Vol. 28, No. 1, Institution of Engineers, Singapore, 1988, pp. 74–83.
15. Olszewski, P., and Y. W. Tan. Traffic in the Central Area—Part II: Speed Characteristics. *IES Journal*, Vol. 31, No. 2, Institution of Engineers, Singapore, 1991, pp. 21–27.

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

Blow Up: Expanding a Complex Random Sample Travel Survey

PETER R. STOPHER AND CHERYL STECHER

In April 1991 the Southern California Association of Governments contracted with the Applied Management and Planning Group to conduct an origin-destination survey of 15,700 households in five Southern California counties. The survey sample was stratified for each county by regional statistical areas (RSAs). Within each county, the sample was stratified also on housing type, vehicle ownership, and household size, yielding a 30-cell sampling matrix for each of the five counties but needing to be reported by the 49 RSAs in the study area. The overall sampling frame was thus a combination of 49 cells (RSAs) and 150 cells (stratification within county). The survey data were collected between April and June 1991. The subsequent file creation and analyses extended far enough into 1992 to enable the use of the 1990 census data to expand the sample to the population. The available census data for expansion consisted of the one-way frequencies of each of the three sociodemographic variables but did not provide any cross tabulations of these variables. The expansion and weighting procedure used to generate population estimates that match as closely as possible the characteristics measured in the 1990 census is described. It is shown that the procedure is relatively simple, even though the sampling procedure was complex. In addition, the sampling errors are reported and are compared with those that would have been obtained from a simple random sample of the entire region.

In April 1991 the Southern California Association of Governments (SCAG) contracted with the Applied Management and Planning Group to conduct an origin-destination survey of 15,700 households in five Southern California counties. The sampling methodology was based on random sampling of households, using random digit dialing (partly screened to eliminate blocks of business numbers and numbers listed in the Yellow Pages). The sampling methodology is based on both geographic and socioeconomic stratification, in the following procedure.

DESCRIPTION OF REGION AND SAMPLING REQUIREMENTS

The Los Angeles region is an urbanized area that covers three entire counties (Los Angeles, Orange, and Ventura), together with the western portion of two counties (San Bernardino and Riverside counties). This urbanized area contained about 14 million residents in 1991. The region extends from the Pacific Ocean on the west to the Mojave Desert on the east, and from San Diego County on the south to the mountain ranges on the north that run from San Bernardino at the eastern end to the coastal ranges at the western end. The area contains a number of major cities, including Los Angeles, Long Beach, Santa Ana, Anaheim, Riverside, San Bernardino, Ventura, and Oxnard. It is estimated that the region contains between 4 million and 5 million households. The majority of the population is located in Los Angeles County (about 8 million people), with the next largest group being in Orange County (about 2 million). The rest are spread almost equally between Ventura, Riverside, and San Bernardino counties.

For planning purposes, the region is divided into regional statistical areas (RSAs), of which there are 44 in urbanized areas and another 12 in rural areas not included as part of the urbanized region (the eastern portions of San Bernardino and Riverside counties and Imperial County to the south and east). The RSAs vary widely in both geographic extent and population, but they are apportioned among the counties in approximate proportion to the populations in each county. RSA boundaries are contiguous with county boundaries and also with 1980 census geography (i.e., each RSA contains whole census tracts and no partial tracts). There are 21 RSAs in Los Angeles County, 10 in Orange County, 6 in Ventura County, 4 in the urban portion of Riverside County, and 3 in the urban portion of San Bernardino County. (The survey area was extended in each of Riverside and San Bernardino counties to include some of the developing western desert areas of Palm Springs and the Coachella Valley, adding three RSAs in San Bernardino and three in Riverside. One RSA in Los Angeles County was deleted from the study because it was predominantly in the mountains and had only 700 listed telephone numbers. The total study area thus covered 49 RSAs.)

Initially, sampling requirements were computed for the region, using Smith's method (*1*) as modified by Stopher (*2*), using a geographic stratification by county and socioeconomic stratification by household size, vehicle ownership, and housing type. For this purpose, household size was divided into five categories, covering each size of household from one person through five persons and up; vehicle ownership was divided into three categories (no vehicle, one vehicle, and two or more vehicles); and housing type was divided into two categories: single-dwelling unit (SDU) and multiple-dwelling unit (MDU). This procedure used the sample standard deviations of trip rates from the 1967 survey but corrected the distribution of households from 1967 to the estimate (at that time) of 1990 population distribution among the categories.

These sampling requirements were based on the objective of achieving a sampling error less than or equal to that of the 1967 survey, with the criterion level being set as ±5 percent

P. R. Stopher, Louisiana Transportation Research Center, 4101 Gourrier Avenue, Baton Rouge, La. 70808. C. Stecher, Applied Management and Planning Group, 12300 Wilshire Boulevard, Suite 430, Los Angeles, Calif. 90025.

error at 95 percent confidence. This led to a sample size in each county that was quite modest, generally falling below that required for statistically accurate modeling of trip distribution and mode choice. The sample sizes also fell below those thought to be politically necessary for the acceptance of population and travel data in such a large urban region.

Through a combination of the political acceptability, available funding for data collection, and the distribution of funding among the counties, it was determined that a total sample of 15,800 households should be targeted for completion.

Sample Design

Initially, it was considered desirable to allocate the sample equally among all RSAs in the targeted area, thus resulting in an initial sample of 320 households from each RSA. This produced targets of 6,400 households in Los Angeles County, 3,200 in Orange County, 2,880 in Riverside County, and 1,600 in each of San Bernardino and Ventura counties. The rationale for equal sample sizes in each RSA was largely that it provides equal sampling errors in each RSA. However, the geographic sampling in this manner also provided a reasonable match between the sources of funding and the sample generated and ensured that the largest samples would be drawn from the most populous areas.

Within each county, it was then desired to ensure that the sample met the minimum requirements of sample size to achieve the maximum error of ± 5 percent at 95 percent confidence for each cell of the three-dimensional socioeconomic matrix. Therefore, the expected sample was distributed among the socioeconomic categories, and the expected sample sizes were compared with those computed from the modified Smith's method. In Los Angeles and Orange counties, it appeared that all cells would exceed significantly the minimum sample size needed. However, it appeared that some cells in the other three counties would contain fewer than the desired number of samples.

The sampling methodology used was to draw a random sample from within each RSA and classify each household during the initial telephone contact into the appropriate housing type, vehicle ownership, and household size cell by county. As sampling approached 90 percent of the target, these totals were examined to determine if any were in danger of not being met. In the event that any cells were less than 90 percent complete while other cells already met or exceeded the sample requirements, remaining samples in that county were shifted from a purely random sampling into a stratified sampling, in which only households from the still-incomplete cells were added to the sample. This was done by using the classifying data in the initial telephone interview as a screening procedure and terminating households that fell into cells that had already reached their target levels. However, of the 16,000 households completing the survey, only 96 were recruited using this targeted approach.

Practical Aspects of Sampling Methodology

It is important to note that the survey mechanism used consisted of an initial telephone interview that classified a household and attempted to recruit the household for diary completion. Then diaries were mailed to all households that agreed to participate, following which the data in the diaries were retrieved through a telephone call. A number of households will refuse to participate at the outset of the first telephone contact, constituting an outright refusal. Others will terminate the interview before recruitment or will refuse to be recruited to complete the diaries. A third group will agree to be recruited but then will not complete diaries or will refuse to furnish the information from the diaries when called. A financial incentive was sent with the mailed-out diaries in order to achieve a higher response rate.

To allow for these various refusals and premature terminations, the initial telephone recruitment seeks to oversample fairly significantly in order to compensate for the losses of sample. The initial recruitment aimed at obtaining approximately 750 households in each RSA and at recruiting in each socioeconomic classification cell about 2.5 times as many households as were considered necessary to meet sampling requirements. However, when the response rates vary by cell from the recruited households (as is always the case), the final sample will be distributed differently from the recruited sample.

EXPANSION METHODOLOGY: SCAG HOUSEHOLD SURVEY

Overview

The methodology to expand the SCAG household survey data consisted of two primary steps: the first was to expand the actual survey responses to represent the total population of households, and the second was to reweight the expanded data to represent the proportion of households by size, housing type, and vehicle ownership. The first of these two steps is based on the number of responding households in each RSA compared with the total number of occupied households residing in the RSA (*3*). (There is a further step to the expansion within this process that accounts for the lack of response by individual household members within any household, where the intent is to expand the number of trips recorded to account for household members that did not complete diaries; however, this step was not performed as part of this study.)

The second step consists of calculating weights from available data to correct for biases in the final samples. Biases arise from two principal sources: (a) self-selection by households concerning response to the survey, wherein experience has shown that households with two or third persons and at least one vehicle are more likely to respond than most other household groupings; and (b) intentional stratified sampling, wherein the sample design was aimed at ensuring certain minimum numbers of households of various subcategories being included in the sample.

Step 1: Expansion

In this step, two expansion activities should be undertaken. The first, which was not done in this study, is at the level of the household and the second, which was done in this study, is at the level of the RSA. For the first step, a trip rate per person would be calculated for each household type, household size, and vehicle ownership level within each county. This per person trip rate would be calculated by dividing the total trips recorded by persons in households of each category and dividing these by the number of diaries from which the sum was obtained. Households would then be identified in which the number of completed diaries retrieved was less than the number of persons reported as living in the household (but not including any persons who completed a diary and returned it indicating that on that day they either did not leave home or were out of town). The person trip rate would be multiplied by the number of individuals that did not respond with a diary in each household category, and this number would be added into the household record.

The second step in the expansion is to compute the ratio of the occupied households in each RSA to the number of responding households in the RSA. This ratio is used as a multiplier for all households in the RSA. For example, the data show that RSA 7 in Los Angeles County had 20,192 occupied households in 1991, and the survey obtained responses from 324 households in that RSA. Therefore, the expansion factor applied to all the households in RSA 7 is 20,192/324, or 62.32. In other words, each surveyed household in RSA 7 represents 62.32 actual households. But even though it appears obvious that this should be the next step in the process, there are good reasons to make this the final step of the expansion process, when the reweighting data are only partially known.

Step 2: Reweighting

Ideally, the second step would be conducted by using three-way cross tabulations of the census data, updated to 1991, showing the actual distributions of households by size, housing type, and vehicle ownership. Although sampling was performed at the county level against these categories, it is more accurate to calculate weights at the RSA level because the bias in the raw survey data stems from both the stratified sampling and the differential response rates that can be expected to differ from RSA to RSA. If the 1990 census totals were available by RSA, then this step is also a simple one, in which the numbers of households of each class as shown by the census for each RSA would be divided by the number of that class found in the sample for each RSA. This factor would be multiplied by the sample number, thus re-creating the actual number of households in each class in each RSA.

Because the 1990 census statistics showing cross tabulations of these three variables were not available in time for use in this expansion effort, an alternative strategy was used. The only census data that were available in the appropriate time frame were the one-way totals of households by housing type, household size, and vehicle ownership (from the May release—STF3). Cross tabulations of any variables were not available.

In the recruitment process for households, five dispositions were possible for each contacted household: the household member who was contacted

1. Refused to answer any questions;
2. Refused to answer questions of household size and automobile ownership, although he or she may have answered prior questions;
3. Refused to participate in the diary survey but did answer the questions on household size and automobile ownership;
4. Answered the telephone interview questions and agreed to participate in the diary survey, from which diaries were not, however, retrieved; or
5. Answered the telephone interview questions and agreed to participate in the diary survey, and diaries were successfully retrieved.

The fifth of these groups is the one that provided the raw survey data. The first two groups provide no usable data. The third and fourth groups provide data from which distributions can be computed of households by household size and vehicle ownership (housing type was not collected from Group 3). Adding Groups 3, 4, and 5 and stratifying by RSA, household size, and vehicle ownership, distributions can be obtained that are less biased than those from the fifth group alone. Determining the proportions of each type of household in each RSA from the summed groups divided by the proportion in the fifth group alone provides a reweighting coefficient that can be applied to each household in each cell of the matrix within each RSA, when census data are lacking. In this project, however, these weights were not used because census data were available in time to be used.

Because census data currently available provide only the one-way distributions on the control variables, an iterative row-and-column balancing (iterative proportional fitting or the Furness method) is used to correct the two-dimensional matrices obtained by taking each of the variables two at a time to create the most probable underlying cross classification, thereby producing weights to redistribute households in each class. In this method, the row and column entries are balanced alternately in iterative steps until the iterations converge to a stable set of cell values that sum to the desired row and column control totals.

In this project, the first reweighting was for dwelling unit type, using the RSA totals of SDUs and MDUs, adjusted for occupied units, as the control. Adjustment factors were obtained from the final iteration and were multiplied through all cells to yield new totals of the sample data, from which automobile ownership statistics were obtained. In the second step, automobile ownership was adjusted through the same procedure by RSA, after which automobile ownership and housing type by county were readjusted to produce county totals for each category that matched the census data. New composite expansion factors were derived from this and applied to the original distribution of households, from which new cell totals were determined and new statistics produced on automobile ownership and housing type.

Again, because the cross-tabulation data were not available, the next item to become available from census releases was that of household size distributions that were provided by SCAG at the RSA level. Using the cell values produced by the preceding factoring step, totals were computed by RSA and county for households in the five household size groups. These were factored to produce the correct RSA totals, following which two successive applications were made of the

Furness method, first to produce household size by automobile ownership and then to produce household size by housing type, after incorporating the adjustments from the previous step.

In the final step, the composite factors were applied to the original cells and the RSA population totals were rebalanced to total RSA expanded population. (It would have had the same effect to have performed all of the preceding adjustment steps on unfactored data, using the proportions of households in each category instead of the absolute numbers, and then factoring the resulting RSA sample populations to the total census populations. Mathematically the results are identical, and there is some appeal in seeing figures throughout the adjustment process that are of the order of magnitude of the total population.)

The procedure described here yielded expanded data that were, for all five counties, less than 1 percent different from the countywide control totals on each of the three categorical variables and that represented an exact match to the census population totals. Figure 1 shows a summary of the steps used to expand the data.

Example

An example of the application of this methodology may be helpful. The example uses one specific RSA to track the effects of the steps in the expansion process. At the time that the expansions were performed, census data were available to provide the 1990 estimates of population, number of households, and distributions of households by vehicle ownership, housing type, and household size. No cross tabulations of these attributes were available, however. Table 1 gives the actual numbers of households recruited that completed diaries for all members of the household present on the day of the survey, by category.

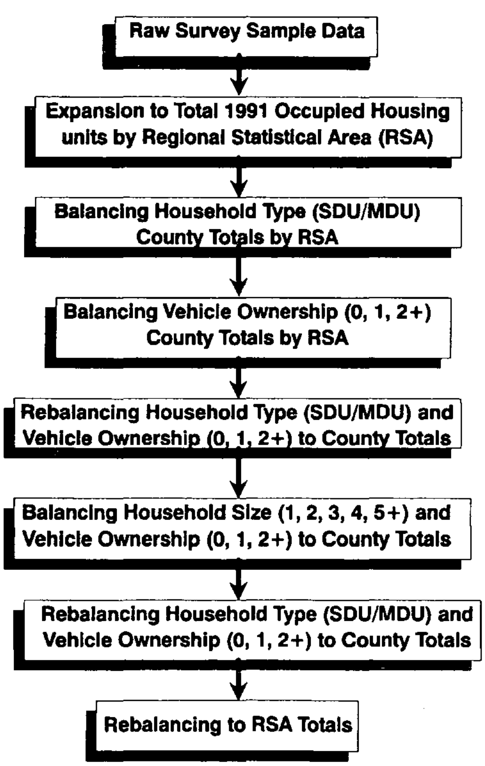

FIGURE 1 Southern California origin-destination survey expansion flow chart.

According to census information, updated to 1991, there were 21,672 housing units in RSA 71, of which 15,909 were SDUs and 5,763 were MDUs. The estimated number of occupied housing units in 1991 was 20,192.

The initial expansion step undertaken was to expand households to the total in each RSA, on the basis of the census figures. The expansion factor for RSA 7 was 62.321 (equal to the ratio of the actual number of occupied housing units, 20,192, divided by the number of completed households, 324). The resulting distribution of households is presented in Table 2.

It should be noted that zero entries in the original table (Table 1) can never change through the expansion process. The next step was to adjust all RSAs by county to the county-level totals of MDUs and SDUs. For Los Angeles County, the total numbers of SDUs and MDUs were 1,516,956 and 1,493,449, respectively. However, the initial expansion of the data, as performed in the step shown in Table 2, generated totals of 1,734,606 SDUs and 1,276,022 MDUs. The Furness method was applied to the RSAs to rebalance to the correct totals, and stable results were obtained after 10 iterations. Applying the resulting adjustment factors produced RSA 7 values given in Table 3.

After this step, census data were obtained on vehicle ownership levels by county. This showed that the expanded data at this point contained 7.05 percent households owning zero vehicles, compared with 11.2 percent in the census; 34.6 percent households owning one vehicle, compared with 35.8 percent in the census; and 58.4 percent households owning two or more vehicles, compared with 53.0 percent in the census. The next step in the procedure was to apply the Furness method by RSA to the vehicle ownership figures. This converged after nine iterations, following which the joint distribution of households by housing type and vehicle ownership for the county was rebalanced to provide jointly the correct totals for each category of housing type and vehicle ownership. The results of this step, which also caused a shift away from the correct totals of households in each RSA, are presented in Table 4.

The next step in the procedure was to balance the county RSAs with the census distribution of household size. Table 5 gives the comparison of the expanded data and the census data, where the difference in the number of households is a result of growth from 1990 to 1991.

Applying the census percentages in Table 5 to the expanded data, new targets were determined for the distribution of households by household size. Two successive applications of the Furness method were then used: in the first, the balancing was done to the household size distribution and the vehicle ownership distribution; in the second, balancing was done to the vehicle ownership and housing type distributions. The results of these steps are presented in Table 6.

One step remained at this point, which was to return the total number of households in the RSA to the original number

TABLE 1 Final Completed Households for RSA 7 by Category

Household Size	Vehicle Ownership						Total
	0		1		2		
	SDU	MDU	SDU	MDU	SDU	MDU	
1	0	1	17	12	6	4	40
2	0	0	6	6	92	36	140
3	0	0	4	3	64	7	78
4	0	0	0	0	42	4	46
5+	0	0	1	0	18	1	20
Total	0	1	28	21	222	52	324

TABLE 2 Initial Expansion Results for RSA 7

Household Size	Vehicle Ownership						Total
	0		1		2		
	SDU	MDU	SDU	MDU	SDU	MDU	
1	0	62	1059	748	374	249	2492
2	0	0	374	374	5734	2244	8726
3	0	0	249	187	3989	436	4861
4	0	0	0	0	2617	249	2866
5+	0	0	62	0	1122	62	1246
Total	0	62	1744	1309	13836	3240	20191

TABLE 3 Second Adjustment of Household Distribution in RSA 7 Using Housing Type

Household Size	Vehicle Ownership						Total
	0		1		2		
	SDU	MDU	SDU	MDU	SDU	MDU	
1	0	78	979	941	345	314	2657
2	0	0	345	470	5296	2822	8933
3	0	0	230	235	3684	549	4698
4	0	0	0	0	2418	314	2732
5+	0	0	58	0	1036	78	1172
Total	0	78	1612	1646	12779	4077	20192

TABLE 4 Rebalanced Distribution for RSA 7 Based on Joint County Distribution of Housing Type and Vehicle Ownership

Household Size	Vehicle Ownership						Total
	0		1		2		
	SDU	MDU	SDU	MDU	SDU	MDU	
1	0	136	1168	1028	346	288	2966
2	0	0	412	514	5307	2589	8822
3	0	0	275	257	3692	504	4728
4	0	0	0	0	2423	288	2711
5+	0	0	69	0	1038	72	1179
Total	0	136	1924	1799	12806	3741	20406

TABLE 5 Comparison of Countywide Household Size Data to Expanded and Adjusted Data at Step 3

Household Size	Expanded Data		Census Data	
	Number	Percent	Number	Percent
1	770,964	25.61	745,661	24.95
2	1,122,512	37.28	835,043	27.94
3	496,478	16.49	474,760	15.89
4	379,740	12.61	417,815	13.98
5+	240,935	8.00	515,148	17.24
Total	3,010,628	100.00	2,988,427	100.00

TABLE 6 Revised Distribution of Households for RSA 7 After Step 5

Household Size	Vehicle Ownership						Total
	0		1		2		
	SDU	MDU	SDU	MDU	SDU	MDU	
1	0	136	1085	1037	311	281	2850
2	0	0	304	412	3791	2009	6516
3	0	0	261	265	3399	504	4429
4	0	0	0	0	2576	332	2908
5+	0	0	146	0	2141	161	2448
Total	0	136	1796	1714	12218	3287	19151

in Table 2. The results of this are given in Table 7, and the final expansion factors are given in Table 8. Table 9 presents a comparison for the entire county of the numbers of households in each category of the three variables and the percentage differences, after applying the final adjusted expansion factors.

Comparing Table 7 with Table 2, it can be seen that the adjustment procedure has made significant changes to the distribution of households by the three categorization variables. As might be expected from an examination of the adjustments required, population has been shifted out of the households that own two or more vehicles (both SDU and MDU) and added into households that own zero or one. In addition, household sizes of one, four, and five-plus have each increased, and the middle two size groups have both decreased. Table 8 indicates that the expansion factors for each cell are quite markedly different, ranging from 41 to 161. This shows the clear need for a more complex expansion procedure than would have occurred using a simple expansion to the RSA population of occupied housing units.

Finally, an examination of Table 9 indicates that the iterative row and column balancing (Furness method) has produced results that are within less than 1 percent error on all county-level demographics. The largest errors occur on house-

TABLE 7 Final Expanded and Adjusted Household Distribution in RSA 7

Household Size	Vehicle Ownership						Total
	0		1		2		
	SDU	MDU	SDU	MDU	SDU	MDU	
1	0	143	1143	1093	327	296	3002
2	0	0	320	434	3997	2118	6869
3	0	0	275	279	3583	531	4668
4	0	0	0	0	2716	350	3066
5+	0	0	153	0	2257	169	2579
Total	0	143	1891	1806	12880	3464	20184

TABLE 8 Final Adjusted Expansion Factors for RSA 7

Household Size	Vehicle Ownership					
	0		1		2	
	SDU	MDU	SDU	MDU	SDU	MDU
1	0	135.7385	63.80844	86.41374	51.87328	70.25033
2	0	0	50.68585	68.64223	41.20522	55.80292
3	0	0	65.3332	88.47868	53.11284	71.92903
4	0	0	0	0	61.32925	83.05626
5+	0	0	146.316	0	118.9481	161.0876

TABLE 9 Comparison of Demographic Distributions After Expansion

Variable	Census	Expanded Sample	Percent Difference
Housing Type			
SDU	1,516,956	1,517,091	0.0089
MDU	1,493,449	1,493,537	0.0059
Vehicle Ownership			
0	337,562	337,559	-0.0009
1	1,078,985	1,078,983	-0.0002
2+	1,594,081	1,594,086	0.0003
Household Size			
1	751,201	757,387	0.8235
2	841,247	844,549	0.3925
3	478,287	476,292	-0.4171
4	420,919	418,549	-0.5631
5+	518,975	513,852	-0.9871

hold size, because that was the variable against which adjustments were made furthest back in the process. With that exception, the county totals are replicated to within less than 1/100 percent.

REFERENCES

1. Smith, M. E. Design of Small-Sample Home-Interview Travel Surveys. In *Transportation Research Record 701*, TRB, National Research Council, Washington, D.C., 1979, pp. 29–35.
2. Stopher, P. R. Small-Sample Home-Interview Travel Surveys: Application and Suggested Modifications. In *Transportation Research Record 886*, TRB, National Research Council, Washington, D.C., 1982, pp. 41–47.
3. Stopher, P. R., and A. H. Meyburg. *Survey Sampling and Multivariate Statistics for Social Scientists and Engineers*. D. C. Heath and Co., Lexington, Mass., 1979, Ch. 4.

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

Factoring Household Travel Surveys

HYUNGJIN KIM, JING LI, STEPHANIE ROODMAN, ASHISH SEN, SIIM SÖÖT, AND ED CHRISTOPHER

Household travel surveys have been conducted recently in a number of metropolitan areas. Metropolitan areas, travel behavior, and data processing capability have changed, requiring current data for analysis and planning; but nonresponse bias persists, often leading to a lack of adequate representation of low-income, low-education, and minority residents. A method to subdivide a study area geographically into microzones to account for the differences in return rates is suggested in conjunction with conventional factoring. In applying this basic step to data collected in suburban Chicago, the representation of families residing in selected areas increased markedly in contrast to factoring by standard techniques, yielding factored data that correspond to county-level data available from standard sources.

Currently there is a surge in collecting and processing travel data in large metropolitan areas. Household travel surveys have been conducted recently in Boston, Baltimore, Dallas, Houston, Cleveland, and Chicago. There are at least three prominent reasons for its occurrence now:

1. Cities continue to decentralize, evolving into urban structures quite unlike those for which data are available. In the past 20 years, for example, the Chicago metropolitan area has had a population increase of only about 4.1 percent, but the amount of land consumed by the urban area has increased by about half (*1*). Nevertheless, many agencies are still using data collected in 1980 or earlier.
2. The 1990 census in general and the work trip information in particular (Census Transportation Planning Package) can be used as a reference for factoring travel surveys conducted now.
3. The advances in computer hardware and software have increased the ability to process, analyze, and display transportation data.

The recent Nationwide Personal Transportation Survey further demonstrates the importance of these data collection efforts by documenting travel growth (*2*). For example, it shows that in the past 21 years (1969–1990), the population of the United States has increased by only 42.2 million inhabitants but the number of licensed drivers has increased by 60.6 million and the number of household vehicles by an astounding 92.7 million. Since much of this growth is in urban areas, the changes in travel demand that these numbers suggest should be examined with fresh data.

H. Kim, J. Li, S. Roodman, A. Sen, and S. Sööt, Urban Transportation Center, MC 357, University of Illinois, Chicago, Ill. 60680. E. Christopher, Chicago Area Transportation Study, 300 West Adams Street, Chicago, Ill. 60606.

The objective of these new travel surveys is that the data reflect the population of the entire study area and their travel patterns. Low response rates among important segments of the population, particularly in minority neighborhoods, can lead to the underrepresentation of these segments in the final data. This in turn leads to their unintended exclusion from the transportation planning process. The purpose of this paper is to suggest a method for improving the factoring process by incorporating the spatial pattern of return rates as an additional step in the factoring process. The aim is to minimize the underrepresentation of population subgroups, regardless of their characteristics.

The survey return rates for suburban Lake County, north of Chicago, used later in the paper to demonstrate the method, significantly correlated with a number of population characteristics. Return rates aggregated by square-mile microzones yield negative correlations (significant at 0.01) with population density, percentage of population with annual incomes under $8,000, percentage of population under age 6, and percentage minority population (Latino, African American, and Asian). College graduation rates were significant at 0.05. Significant positive correlations were recorded with percentage of households with incomes exceeding $50,000 and percentage of households living in single-family homes.

This implies that higher response rates are expected from affluent, low-density suburban communities than those from high-density urban areas with lower incomes. It is reasonable to assume that the lower return rates are found in areas where mobility is limited. The interest in the mobility-disadvantaged makes it necessary to ensure that they are represented adequately in the survey data. It is, however, impractical to try to factor the data for all of the variables described earlier. Therefore, the authors suggest treating the problem directly, using the return rates. Furthermore, this method does not require the respondents to identify what many individuals consider to be confidential information.

Transportation surveying efforts usually include factoring by household demographics, such as household size, number of workers, and automobile availability, which are commonly held to account for variations in travel behavior. Although these demographics are useful for providing weights, they do not always account adequately for minority populations, which may have the same household size and automobile availability but different return rates and, more important, different travel patterns. The method described here addresses this problem directly. It is illustrated in a later section by using the 1989 Lake County portion of the Chicago Area Transportation Study Household Travel Survey.

It should be mentioned that the importance of factoring stems from the use of these data for descriptive purposes.

When the data are used for model development it may not be necessary to factor, and in some circumstances it may be ill-advised (see the paper by Thakuriah et al. in this Record).

BACKGROUND: BASICS OF FACTORING

This paper focuses on mail surveys, but the problems identified and the method suggested also apply to telephone surveying and personal interviewing. Telephone surveying has recently become popular, but research continues on ways to improve mail surveys as an effective way to collect data (*3,4*). Cost-effectiveness and steady return rates are the principle advantages of mail surveys (*5*).

Work has also focused on methods of factoring survey data. FHWA has provided guidelines on ways to factor surveys (*6*), and alternative procedures have been evaluated (*7*). Applications have been many: the Metropolitan Transportation Commission study (MTC) (*8*) and the more recent study in Phoenix (*9*) both illustrate the effective use of zone-specific demographic data in factoring.

The factoring method proposed here consists of three steps. First the zonal structure is selected, then demographic data are obtained and tabulated, and finally the weights are computed. A fourth step, establishing reasonableness, is advisable but not pertinent.

Factoring Zones

The selection of zones is a classic problem. It is a trade-off between using small zones to ensure internal homogeneity and using large zones so that data handling costs are minimized. Zonal homogeneity is important and can be measured by variables such as housing costs, distribution of jobs, proximity to amenities, and municipal zoning ordinances. Ideally, transportation planning zones would follow these neighborhood differences and the variation inherent in the corresponding travel behavior.

Most agencies use zonal geography to account at least partially for these patterns. The studies of MTC (San Francisco Bay area) and Phoenix used 34 and 8 superdistricts, respectively. It is unlikely, however, that factoring-zone boundaries follow the socioeconomically defined neighborhoods. Some demographic communities are split or become minority sections within factoring zones. As a consequence, residents may be grouped with individuals unlike themselves.

The authors propose the following two-stage procedure to address this problem. First, a two-tier zonal system is identified that includes basic factoring zones or districts and a system of subzones, here called microzones. In the Lake County application, the basic factoring zones are townships in the township and range system, and the microzones are 1-mi^2 neighborhoods. Since the typical township is 6 mi^2, there are 36 microzones in a basic factoring zone (Figure 1). The microzones do not need to be this small, but they should be easily identifiable zones that are internally homogeneous. In the Chicago area the township and range geography is widely used for planning purposes because major arterials often constitute the boundaries and therefore they are easily identified. Moreover, these arterials frequently are the delimiters of socioeconomic differences between adjacent neighborhoods.

Second, the microzone return rates are computed and mapped. If the basic factoring zone has relatively high and uniform microzone return rates, it remains unaltered. For example, in the southern quarter of Lake County almost all the microzones had adequate return rates (more than 20 percent) or fewer than 10 survey instruments mailed and therefore the return rates were not considered. Near the northwestern and northeastern portions of the county are areas with low return rates. Here microzones with particularly low return rates are combined into a factoring subzone, whether they are contiguous or not. The rest of the township becomes another factoring subzone. If further discrepancies in return rates remain, additional factoring subzones may be created and the original zone may be split into any number of subzones. In this way the original number of factoring zones increases, but the increase may well be moderate.

Providing special treatment to microzone areas within the basic factoring zone with particularly low response rates creates a means of reasonable representation for groups that might otherwise be underrepresented, even with the demographic factoring common to most studies.

Advances in address matching and geographic information systems make this procedure feasible. A requisite is the ability to build different geographically defined zones (microzones and basic factoring zones). Local convention may well dictate whether microzones are square miles, census tracts, or other zones.

Present computer technology allows the use of a large number of analysis and factoring zones, but there are still several practical reasons to keep the number moderate. First, communicating information in reports would be difficult if tables had hundreds of zones. Not only is the amount of information overwhelming, but as the number of zones increases, the likelihood decreases that readers understand where these zones are located. Second, for longitudinal studies it would be useful to maintain consistency in zonal geography. Therefore, the increase in the number of zones should be controlled.

Factoring on Basis of Demographics

In the second step, the data are factored using traditional demographic characteristics that account for much of the variation in trip-making behavior. In the authors' survey of large metropolitan transportation planning agencies across the country, 11 of the 23 organizations surveyed used some method of factoring in their survey work. Of these, the authors received information on eight, and six crossed specific variables and used direct extrapolation in the matrix (two were unsure of the factoring method) using census information as their standard. The most commonly used variables—household size and number of vehicles per household—were used by half of the organizations, and they are used here. Other variables such as number of workers in the household would serve well as alternatives.

Computing Weights

The third and final step is the computation of the weights for factoring. This is discussed later.

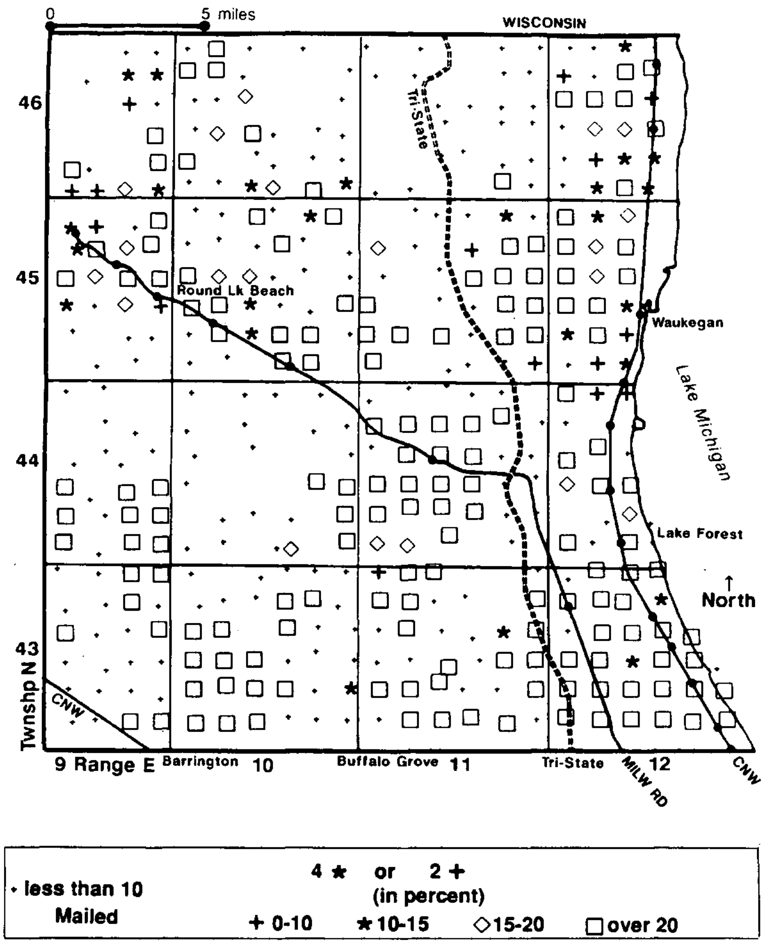

FIGURE 1 **Lake County return rate by square-mile zones.**

FACTORING: PROCEDURE

In this section we will describe the data and the computational steps to determine the factoring weights. A sample application illustrating the procedure is provided later.

Input Data

For the purpose of factoring there are four key pieces of information in a mail-out/mail-back survey:

M_c = number of questionnaires mailed to Zone c ($c \in b$, and $b \in a$),

Q_c = number of questionnaires returned from Zone c ($c \in b$, and $b \in a$),

S_{ad} = number of survey households in Zone a that belong to Demographic Category d, and

H_{ad} = number of households reported by census in Zone a that belong to Demographic Category d,

where

c = microzone (e.g., 1-mi² zone),
b = combination of microzones or factoring subzone, and
a = large zone or basic factoring zone (e.g., a township).

Step 1: Identifying Different Factoring Areas Using Return Rates

The intent of the first step is to identify areas with approximately the same return rates. These are either basic factoring zones a or, if they are not satisfactory, a combination of the microzones (sum of c zones).

Figure 1 illustrates a sample case in which there are initially 16 basic factoring zones (townships). The return rate, R_c, of each microzone c,

$$R_c = Q_c/M_c \tag{1}$$

is used to form subzones (b's) with approximately the same return rates. Along the southern tier the townships are found to be homogeneous and therefore remain unaltered. The two townships in the northeastern corner of Lake County along Lake Michigan (12-45 and 12-46) have many microzones with very low return rates. These microzones, regardless of contiguity, are aggregated as a separate factoring zone, yielding two factoring zones in the township. Although in this paper the aggregated area is called a subzone, it is a factoring zone just like the unaltered township.

Step 2: Preparing Demographic Data—Household Size and Vehicle Availability

After defining subzones on the basis of return rates in Step 1, the demographic data for factoring need to be prepared. As stated, household size and vehicle availability are used. Since low numbers of households in each cell of the cross tabulation in Figure 2 would result in high variances, the authors combined cells to ensure minimum sample sizes in each factoring cell. For example, many single-person households with more than one vehicle would not be expected, so these are combined.

In Lake County the six aggregates of Cells I through VI were selected as shown in Figure 2. For example, Category I includes all zero-vehicle households and those one-vehicle households with three or more members. By tallying the data from DuPage and Cook counties (the two most populated counties in metropolitan Chicago) by the six categories in each township, it is found that in order of frequency they are V, IV, VI, I, II, and III. Although on average Category III was the least populated, there was at least one township in Lake County in which it recorded the largest number of households.

Step 3: Computing Factoring Weights

For factoring zones that are not subdivided, such as those in the southern portion of Lake County (Figure 1), factoring is simple and direct. Each record is weighted by

$$W_{ad} = H_{ad}/S_{ad} \tag{2}$$

And if the cross tabulation in Figure 3 can be completed for each microzone (c's), and thereby for the aggregates of these microzones (b's) then Equation 2 may be used.

In cases in which there are several factoring subzones within a township and the cross tabulation data are not available, the calculations deriving the weights are described here.

Determine the weight, W_{bd}, such that (see Figure 3)

$$\sum_{b=1}^{n} W_{bd}S_{bd} = H_{+d}$$

and

$$\sum_{d=1}^{m} W_{bd}S_{bd} = H_{b+} \tag{3}$$

The example in Figure 1 has only two subzones in each basic factoring zone (township).

Household Size

Vehicles Available	1	2	3	4+
0	I			
1	II	III		
2		IV	V	
3+	VI			

FIGURE 2 Cross tabulation of data for each factoring area.

Household size and Vehicles available (d)

Subzone (b)*	I	II	III	IV	V	VI	H_{b+}
1							
2							
·							
n							
H_{+d}							

FIGURE 3 Demographic cross tabulation for each large zone.

Since detailed census data are not yet available, the overall weights, W_b and W_d, for each subzone and demographic category can be used in Equation 3. Although there are two subzones in each basic factoring zone in Figure 1, the notation of Equation 3 illustrates n subzones.

The weights

$$W_b = H_{b+}/S_{b+} \tag{4}$$

are applied to S_{bd} to satisfy the second part of Equation 3. However, the first part of Equation 3 is not satisfied. In other words, W_b scales up the number of households in each Zone b to match with census data, but it might not match with the census demographic categories. Therefore, adjusting weights are needed. The adjusting weight is

$$W_d = H_{+d} \Big/ \sum_{b=1}^{n} W_b S_{bd} \tag{5}$$

It is verified that $W_b W_d S_{bd}$ satisfies the first part of Equation 3 but not necessarily the second. To satisfy the second part, continue this adjusting process on each occasion satisfying one of the two parts of Equation 3. After r pairs of such steps, compute the general form of Equations 3 and 4 as

$$W_b^{r+1} = \frac{H_{b+}}{\Sigma_{d=1}^{m} \left\{ \prod_{p=1}^{r} W_b^p W_d^p \right\} S_{bd}} \tag{6}$$

and

$$W_d^{r+1} = \frac{H_{+d}}{\Sigma_{b=1}^{n} \left\{ \prod_{p=1}^{r} W_b^p W_d^p \right\} W_b^{r+1} S_{bd}} \tag{7}$$

This is the Deming-Stephan or Furness (DSF) procedure. It is easily generalized to more dimensions—for example, if there are more demographic categories or more zones.

Final Factor

The final factor, W_{bd}, can be calculated using W_b and W_d obtained from the DSF procedure (Equations 6 and 7). The final factor is

$$\left\{ \prod_{p=1}^{r} W_b^p W_d^p \right\} \rightarrow W_{bd} \tag{8}$$

In the example presented later, W_{bd} converges in five iterations.

APPLICATION

The proposed method of factoring was applied using the Chicago Area Transportation Study Household Travel Survey. The data were collected from Lake County, Illinois, a county of 173,996 households in 1990. It is a rapidly growing county in northern suburban Chicago, bordering Wisconsin, Lake Michigan, and Cook County (Chicago). Demographically, the county has a large range of characteristics, with wealthy neighborhoods encompassing large estates as well as sizable low-income communities.

The three principal minority groups—African Americans, Latinos, and Asians and Pacific Islanders—constitute about 16 percent of the population, with Latinos surpassing African Americans in the 1980s as the largest group. In sum, the county is a low-density suburban area with great contrasts.

Household Survey

Survey instruments were mailed to 9,143 households on the basis of square-mile zones of residence, or the microzones. Since there were approximately 500 such microzones in the county, and this county constituted only one of seven counties in the entire study, it was not practical to make the microzone the basic factoring zone. Instead 16 townships, 36 mi^2 each, constituted the basic factoring zones.

The average return rate in the county was 27.2 percent, but the range per microzone was more than 50 percentage points. The overall pattern by microzone was considered, and zones with fewer than 15 percent, approximately half of the county average, were flagged. If there were enough microzones in a township, they were combined into a separate factoring subzone in each township. Six townships had sufficient numbers of microzones with low return rates that they were combined into factoring subzones. For the 10 townships not subdivided, the response rate ranged from 25 to 37 percent. In sum there were 22 factoring zones.

Results of Proposed Method

The proposed method was used, and a summary of the results for the six townships that were subdivided is presented in Table 1. It indicates, for example, that Township 9-46 (in the very northwestern corner of the county) was one of six in which two subzones were defined. Considering the subzone with a low return rate, if the survey returns were factored only by the six demographic categories of Figure 2, then this portion of Township 9-46 (the subzone) would be factored to 1,285 households. Using the suggested method, this area of low return rate is factored to 2,454 households, or an increase of 1,169 households. In this instance factoring by only demographic characteristics would have vastly underrepresented the travel generated by the subzone with low return rates. In essence, what has been achieved is a geographic redistribution of survey weights within the township without changing the township total.

TABLE 1 Effect of Proposed Method

	Number of Households in Selected Low Return-Rate Subzones		
Township[a]	Proposed Method	Conventional Method	Township
9-46	2,454	1,285	4,389
12-46	4,447	2,766	12,325
9-45	2,070	860	6,709
12-44	1,533	283	7,840
11-45	1,106	253	12,440
12-45	12,018	6,552	29,097

[a]See Figure 1 for Cartesian reference system used to identify townships.

There is little doubt from Table 1 that the proposed method has an effect on the factoring results. For the conventional method it is likely that reasonableness checks would alter the results and bring them closer to the actual. Still, there is such a sizable discrepancy between the two results in Table 1 that many reasonableness checks and adjustments would probably be necessary for the conventional method. Each adjustment may also bring undesirable effects.

REASONABLENESS CHECKS

This survey has three components: the household, person, and trip files. A systematic check of the factored data would include an examination of all three files. Some individuals in a household may not have been included, and certainly not all trips were included for all respondents. Files that do not yield adequate results can be given new weights to compensate for missing data. Ideally, the weights computed for the household file are adequate for the person and trip files, and additional weights need not be computed.

The data structure of the Nationwide Personal Transportation Survey includes a half-dozen separate files, each with its own factors. Although this approach achieves the desired effects on a file-by-file basis, the data that overlap between files do not match, highlighting the apparent inconsistencies in the data. For example, the number of vehicles is reported in several files but the household and vehicles files yield different totals. This is unavoidable when each file carries its own weight.

Conversely, if only the household file is factored there is the risk that the data in the person and the trip files are not properly adjusted. It is essential that great care be taken to factor the household data properly so that the person and trip files need not be factored separately.

After the household file was factored and these factors were applied to the person and trip files, checks of reasonableness were performed on a selected variables from each of the two files. Total population was selected to check the person file, and commuter rail use and miles traveled were used for the trip file. Since the factoring procedure was based on household size and number of workers per household, it was not necessary to check the household files.

A tally of the number of persons yields a population total that is 98 percent of the 1990 census figure for Lake County, indicating that the person file is adjusted adequately. A comparison of commuter rail users, a test of the trip file, is slightly more difficult since ridership data are reported by station whereas survey data are aggregated by township. Station data within a township can be added, but their service areas typically do not follow township boundaries. Table 2 summarizes the commuter rail ridership data for selected townships, where service area definition appeared obvious.

Since the survey data overstate the number of actual commuter rail users, as one would expect from a survey that has higher returns from high-income neighborhoods, the greater emphasis on low-return-rate zones has not overadjusted for this group. In sum, the survey data match the actual ridership data well enough that recomputation of the weights does not appear to be necessary.

The survey trip file also reports trip origins and destinations, coded by ½-mi^2 zones. Summing the air-line distances provides an estimate of 9 million mi of travel for Lake County. Converting these miles to route miles of actual travel adds about 20 percent, yielding an estimate of 11 million mi.

Finally, since commercial traffic is not included in the survey, the estimate of a million such miles needs to be added, resulting in a total estimate of 12 million vehicle miles traveled. This compares very favorably with the estimate of 11.9 million mi from the State of Illinois Roadway file for Lake County (1989) and the CATS assignment model estimate of 12.1 million mi, suggesting that no further adjustment of the trip file is necessary.

If the data do not match the expectation, they need to be adjusted accordingly, using the DSF procedure described earlier. Clearly, the problems associated with readjusting the weights can best be diminished by starting with data that are well factored. Adjustments of the weights are occasionally necessary, but this process cannot create data that do not exist in the survey.

CONCLUSIONS

Considerable research has been conducted on alleviating and controlling for bias in travel surveys. Many agencies use a variation of the factoring method based on demographic characteristics. Although these methods have considerable merit, they do not necessarily address a major deficiency in many travel surveys—namely, that key population subgroups characterized by low incomes and low education achievement rates tend to remain underrepresented. When income information is not solicited it cannot be used to adjust the results, nor should it be used when the information is solicited but the responses are unreliable. It is for these cases that this method is useful.

Since it is likely that the underrepresented population is mobility-disadvantaged, it is particularly important that they be included in a transportation data base. In suburban Chicago the proposed method increased—in contrast to conventional factoring—the representation in low-return-rate areas by more than 10,000 households, yet it did not overrepresent this group.

Two advantages of the method are its simplicity and cost. It is logical and easily applied, especially given the ability to manipulate geographic information by computer.

TABLE 2 Commuter Rail Use: Survey Estimates and Metra Station Data Aggregated by Township

Township	Metra Data	CATS Survey	Difference
12-46	197	211	14
9-45	1,168	1,068	80
11-44	2,331	2,141	190
10-45	1,339	1,203[a]	136
Total	17,604	18,039[b]	435

[a]Includes adjacent townships without Metra stations.
[b]Assumes that 5,000 Lake County Metra commuter rail users crossed county line to closest station.

REFERENCES

1. Gridlock Without Growth? *Planning in Northern Illinois*, Vol. 31, No. 1, Summer 1991.
2. Hu, P., and J. Young. *Summary of Travel Trends, 1990 National Personal Transportation Survey*. FHWA, U.S. Department of Transportation, March 1992.
3. Ayidiya, S., and M. McClendon. Response Effects in Mail Surveys. *Public Opinion Quarterly*, Vol. 54, No. 2, 1990, pp. 229–247.
4. Bishop, G., H. Hippler, N. Schwartz, and F. Strack. A Comparison of Response Effects in Self-Administered Surveys. In *Telephone Survey Methodology* (R. Graves, ed.), Wiley, New York, 1988, pp. 321–340.
5. Harvey, L. Factors Affecting Response Rates of Mailed Questionnaires: A Comprehensive Review. *Journal of the Market Research Society*, Vol. 29, No. 3, 1987, pp. 341–353.
6. *Urban Origin-Destination Surveys*. FHWA, U.S. Department of Transportation, 1973, Ch. 7.
7. Ollman, L. T., S. M. Howe, K. W. Kloeber, and G. S. Cohen. Marginal Weighting of Transportation Survey Data. In *Transportation Research Record 677*, TRB, National Research Council, Washington, D.C., 1979, pp. 73–76.
8. *Sample Weighting and Trip Expansion*. Working Paper 4A, 1981 MTC Travel Survey. Metropolitan Transportation Commission, Berkeley, Calif., 1983.
9. *1988/1989 Household Travel Survey*. Behavior Research Center, Inc., Phoenix, Ariz., 1989.

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

Using Machine Vision (Video Imaging) Technology To Collect Transportation Data

MICHAEL KYTE, ASHFAQ KHAN, AND KISHORE KAGOLANU

The results of a series of tests of machine vision, or video imaging, technology in the measurement of various kinds of transportation data are presented. The Autoscope system is used to collect flow rate data at both interrupted and uninterrupted flow facilities and vehicle size data at a port of entry. The data are compared with manually measured data. The results show that machine vision technology offers excellent potential for effectively and efficiently collecting transportation data, particularly when combined with postprocessors that can identify special data characteristics such as vehicle paths or vehicle length or height.

A variety of increasingly sophisticated systems have been developed over the years to monitor and control transportation facility operations. Signalized intersections are now controlled using electronics to monitor vehicle arrivals and departures and to implement optimal timing schemes to minimize delays and encourage vehicle progression. Freeways are monitored and controlled using vehicle detection algorithms designed to identify congestion and other incidents that cause delay.

A new generation of technology based on video imaging and machine vision concepts is being developed to improve the accuracy of measuring traffic flow conditions, to reduce the response time to changes in traffic conditions, and to eliminate the persistent maintenance problems that are common to inductive loop in-pavement systems. One such technology, the Autoscope system, was developed by the University of Minnesota and Image Sensing Systems, Inc., and has been described by Michalopoulos et al. (*1*). The University of Idaho has been using the Autoscope 2002 system since 1991 to develop methods for automatically collecting traffic flow and delay data at intersections, to monitor freeway traffic operations, and to measure vehicle lengths and heights at ports of entry. Machine vision technology automates these processes by extracting traffic flow and vehicle size data directly from video images.

Machine vision technology combines video and computer technology to extract traffic data from video images. Video images (either live through an on-line video camera or on tape) are taken of traffic flowing through an intersection or along an arterial or highway. The images are transmitted from a video playback machine into a specially equipped personal computer. The personal computer, with a frame grabber board, takes the video image and digitizes and stores each image at a rate of 30 frames per second. A second computer, the Autoscope, processes each digitized image. The Autoscope recognizes the video content of the stationary background image and identifies any changes in the luminance as potential vehicle movements. Special algorithms sort out actual vehicle movements from other spurious or nonrelevant image changes. The user identifies detector locations (points at which he or she wants to measure traffic movements) on the video display using a mouse. The computer identifies all vehicle movements passing detector points that have been identified by the user. The computer stores the time that each detector is activated by a vehicle (i.e. when the presence of the vehicle is detected) and the duration of the presence detection.

To measure turning movement flow rates, however, two other factors must be considered. First, additional detectors must be included that correspond to points along a path of travel through the intersection; second, the detector activation times must be considered together so that possible vehicle movements or paths can be identified. Figure 1 shows an example of the data produced by Autoscope that can be used to identify a turning movement. Detector 0 is activated at 8:00:03.253, and Detector 6 is activated 2.101 sec later. The event described by these consecutive detector activations is almost certainly a northbound left-turning movement, in the absence of other detector data. If other detectors have been activated during this time, however, the logic imbedded in the software must be able to identify this turning movement event out of other possible events.

This postprocessing of the Autoscope data can be used to generate more complex information about the traffic flow characteristics by combining data from two or more detectors. Postprocessors can be developed that can produce, in addition to turning movement data, vehicle delay and queue lengths at an intersection, vehicle speeds along an arterial or freeway, or vehicle length, width, and height data at a port of entry. The common link in determining each of these data is the identification of a specified event by the analysis of patterns or sequences in the raw Autoscope detector data. For example, the simultaneous activation of seven detectors, with a known geometric relationship, might indicate a truck that is between 3.8 and 3.9 m in height. Or the timing of the consecutive activation of two detectors at a known distance apart might indicate a vehicle traveling at 45 km/hr. Once an event is characterized by a given sequence or pattern of detectors (and suitable time delays or constraints are accounted for), it can be identified by an organized search routine or algorithm applied to the raw Autoscope detector data. The general principle is this: the data from a given detector activation pattern is used to determine that the event has occurred.

Civil Engineering Department, University of Idaho, Moscow, Idaho 83843.

The purpose of this paper is to report the results of a research project on the application of the Autoscope 2002 system to the measurement of three broad classes of transportation data: uninterrupted traffic flow, interrupted traffic flow, and vehicle size. For each class of data, the following information is presented: (a) objective of each measurement test, (b) definition of variable(s) to be measured, (c) description of measurement methodology, (d) presentation of results, and (e) discussion of results and summary of relevant conclusions.

BACKGROUND

Video Imaging Systems

One of the first applications of video imaging technology in transportation was in 1978 when FHWA contracted with the Jet Propulsion Laboratory (JPL) to investigate the feasibility of using video sensors and image processing for automatic traffic monitoring. This study culminated in the development and testing of a breadboard system, the Wide Area Detection System (WADS), for vehicle detection and vehicle velocity measurement (*2*). Two studies were completed recently to evaluate currently available video imaging technologies for highway operations. One study, at the University of California at Berkeley, described the components, advantages, and disadvantages of image processing technology for traffic data collection (*2*). The second study, conducted for the California Department of Transportation by the California Polytechnic State University, evaluated the performance characteristics of several image processing technologies (*3*). This study identified eight systems having good potential for analyzing traffic data. These systems include the Aspex Traffic Analysis System (ATAS); the Camera and Computer Aided Traffic Sensor (CCATS) by Devlonics in Belgium; Sigru, developed by Eliop in Spain; the Traffic Analysis System (TAS); Titan, a French system under development by the Institut National de Recherche sur les Transports et leur Securité; Traffic Tracker; Tulip; and Autoscope. The study considered the performance of each system in measuring different traffic parameters under a variety of site conditions including congested flow, fog, snow, rain, nighttime, and poor lighting conditions.

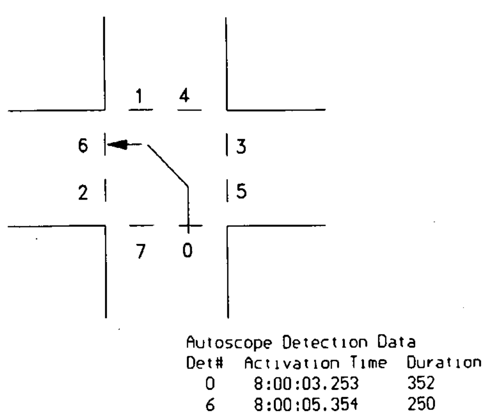

FIGURE 1 Sample Autoscope detector layout and data file for turning movement data.

Applications of Video Imaging

Several recent research projects have tested the use of video imaging technology for the collection of transportation data. At the Environmental Research Institute of Michigan (ERIM), Gilbert and Holmes developed real-time image processing algorithms for detecting and tracking vehicles in actual traffic settings (*4*). The principle that they used for vehicle detection was based on the comparison of the difference in illumination between a defined region of a scene and a passing vehicle. The background color of the scene is taken as the standard of reference, and any passing object with a different light intensity or color is detected as a vehicle.

Frame differencing has been used successfully in two studies. Vieren et al. used frame differencing to define the edges of moving vehicles in successive frames (*5*). Sethi and Brillhart employed video imaging technology to detect the nonconforming behavior of vehicles on roadways in addition to providing normal traffic statistics for traffic monitoring (*6*). They used frame differencing to identify centroids of vehicles for tracking.

Troutbeck and Dods used the Video Detection Data Acquisition System (VADAS) with video data collected from cameras mounted on a 10-m-high telescopic mast to determine vehicle paths (*7*). Taylor et al. also used VADAS for headway and speed data collection on freeways and parking lots (*8*).

The University of Minnesota studied the performance of the Autoscope system for incident detection on a freeway (*9*). In this study, a 5.6-km machine vision laboratory with 38 cameras was designed on Interstate 394 for deployment and validation of the incident detection system. The study reported a false alarm rate of only 3 percent.

The application of video imaging has been used to evaluate pavement surface distress such as cracks, potholes, and depressions. As part of NCHRP Project 1-27, a system was developed to process video images to identify, quantify, and classify pavement distress in terms of types, severity, and extent (*10*). The results have shown an error of 5 to 17 percent, and in some instances it proved to be superior to the visual inspection method currently practiced by the personnel in this field.

Koutsopoulos and El Sanhouri developed a method for automatic interpretation of asphalt pavement distresses, recorded on video or photographic film, with emphasis on segmentation and classification of digitized distress pavement images (*11*). Segmentation is the process of extracting objects of interest from the background, in this case the pavement distresses such as cracks and potholes, and classification is the process of identification of distress type. The preliminary results from this study indicate that it is feasible to automate the process of pavement image analysis.

Lu et al. have illustrated the application of image processing for measuring pedestrian volumes at intersections (*12*). They have developed an algorithm based on television image sequence analysis that automatically counts the number of pedestrians on crosswalks in daytime periods and automatically determines their volume. The tripwire method of detection is employed for this purpose. A band of pixels on the TV image corresponding to the crosswalk are placed perpendicular to the direction of pedestrian movement. A 94 percent accuracy rate was reported in the pedestrian volume measurements.

MEASUREMENT OF UNINTERRUPTED TRAFFIC FLOW

One of the original motivations for developing the Autoscope system was to provide a more comprehensive method of monitoring and controlling freeway traffic operations. Inductive loop systems can provide the flow, speed, and occupancy data required by standard freeway traffic management systems, but machine vision technology offers two primary advantages over inductive loop technology: (a) the concern with maintenance of in-pavement loop detectors is eliminated, and (b) the additional visual information available to traffic engineers monitoring the flow of traffic enables better decisions in incident response.

Objective of Measurement Test and Variable To Be Measured

The objective of measuring uninterrupted traffic flow is to determine how well the Autoscope system can measure freeway traffic flow rates as compared with manual counts measured by an observer. In particular, the relative effectiveness of various detector layout configurations and camera angles was investigated.

The variable to be measured in this test is vehicle volume, in vehicles per time period.

Description of Measurement Methodology

Video cameras were located on overpasses at two locations on I-84 in Boise, Idaho, and videotapes were made for five time periods. Traffic was traveling away from the camera in three of the scenes and toward the camera in two of the scenes. Both horizontally and vertically oriented detectors were tested. (Horizontally oriented detectors are oriented perpendicular to the traffic flow, and vertically oriented detectors are oriented parallel to the flow.) Autoscope was used to collect flow rate data. To determine the accuracy of the Autoscope system, comparison counts were made using a personal computer and the Traffic Data Input Program (TDIP) (*13*). TDIP requires the user to press a key on the computer keyboard corresponding to a particular vehicle event, such as a vehicle passing a point in a given traffic lane. Like Autoscope, TDIP produces a data file of vehicle passage times for each event and summaries of flow rates for each movement that is monitored.

Results

The results of the volume counts from both Autoscope and TDIP are given in Table 1. The table includes 13 comparisons and shows the direction of the traffic flow with respect to the

TABLE 1 Uninterrupted Flow Data

Scene	Duration (minutes)	Lane	Direction	Detector Orientation	Volumes		Percent Deviation	MAPD 5-Minute Data
					Autoscope	TDIP		
2	20	1	Away	H	177	174	-1.7	9.0
2	20	2	Away	H	177	162	-9.3	11.8
2	20	3	Away	V	132	174	24.1	24.5
3	27	1	Toward	H	179	206	13.1	15.9
3	27	2	Toward	H	218	230	5.2	5.2
4	60	1	Away	H	820	825	0.6	2.4
4	60	2	Away	H	424	431	1.6	4.2
4	60	3	Away	H	872	939	7.1	7.5
5	60	1	Away	H	1306	1325	1.4	2.7
5	60	2	Away	H	399	419	4.8	6.3
5	60	3	Away	H	1542	1627	5.2	5.1
7	60	1	Toward	H	1070	962	-11.2	18.9
7	60	2	Toward	H	1370	1346	-1.8	13.2

Notes:

Since neither the Autoscope data nor the TDIP data can be assumed to be completely accurate, percent errors cannot be calculated. Thus in the comparisons presented in this and subsequent tables, the terms *deviation* or *correspondence* are used.
The *Percent Deviation* is calculated with respect to the TDIP data.
The *MAPD 5-Minute Data* is the mean absolute percent deviation calculated between the TDIP and Autoscope data for each 5-minute period for which data are available.
***Direction* indicates whether traffic is traveling away from or toward the camera.**
***Detector orientation* indicates the orientation of the video detector with respect to the flow of traffic. Horizontal orientation (H) indicates that the detector is perpendicular to the flow of traffic; vertical orientation (V) indicates that the detector is parallel to the flow of traffic.**

camera, the orientation of the video detector, and the volumes collected from Autoscope and from TDIP.

Three categories should be noted. The first category shows the deviation calculations for the scenes in which traffic is moving away from the camera and the detector orientation is horizontal. The percentage deviation for these eight comparisons ranged from 0.6 to 9.3 percent, with a mean of 4.0 percent. The second category shows the deviation calculations for the scenes in which traffic is moving toward the camera and the detector orientation is horizontal. In this case, the percentage deviations were somewhat higher, ranging from 1.8 to 13.1 percent. The third category shows the deviation for one scene in which traffic is moving away from the camera and the detector orientation is vertical. This configuration showed the highest percentage of deviation, 24.1 percent.

Discussion of Results

Autoscope clearly has the potential to produce accurate volume counts as shown by the high level of correspondence with the manual volume counts, though care and experience are both needed in the development of optimal video detector layouts and camera placement. The most accurate measurements were obtained when the video camera was placed such that traffic was moving away from the camera and horizontal video detectors were used to collect the volume data. For this configuration, the mean deviation between Autoscope and TDIP volume counts was 4 percent. Mean absolute deviation for 5-min volume counts ranged from 2.4 to 11.8 percent.

Less accurate measurements resulted when traffic was moving toward the camera or when vertical detectors were used. Accuracy was also reduced when daytime light levels were low, such as in the early morning hours, as was the case for Scene 7 in Table 1.

MEASUREMENT OF INTERRUPTED TRAFFIC FLOW

Autoscope was also developed to detect vehicle presence at signalized intersections as a replacement for inductive in-pavement loop detectors. In this set of tests, machine vision technology is used to measure both approach and turning movement flow rates at an intersection. Although the technique used to measure approach flow rate is the same as reported for the freeway volume counts earlier, the turning movement counts involve the development of a postprocessor algorithm.

Objective of Measurement Test and Variable To Be Measured

The objective of these tests is to determine how well the Autoscope system can measure approach volumes and turning movement volumes as compared with manual counts measured by an observer using TDIP.

The variable to be measured in these tests is vehicle volume counts.

Description of Measurement Methodology

Video cameras were placed at elevated locations near four stop-controlled intersections in Pullman, Washington, and Moscow, Idaho. In each case, a single video camera was used to record traffic movements through the entire intersection. Videotapes were made for 12 time periods. Video detectors were located on each approach; to estimate turning movement volumes, video detectors were also placed on the departure lane for each direction, as shown in Figure 2. An algorithm was developed to match entry and exit detectors that fit within a preset time range and that represented logical vehicle paths. The application of this algorithm to the raw Autoscope detector data yielded an estimate of turning movement volumes.

To determine the accuracy of the Autoscope measurements and the turning movement postprocessor, TDIP was used to develop a corresponding set of approach volumes and turning movement counts.

Results

Although a freeway presents a relatively uncluttered traffic scene, an intersection scene can include a significant amount of potentially extraneous images, such as pedestrians. In addition, vehicles do not always follow well-defined paths as they travel through an intersection. For this reason, the placement of the camera is crucial in obtaining a clear and unobstructed view of the vehicles. The height of the camera and angle of the camera view determine whether the image of each vehicle remains unoccluded by either other vehicles or other items that appear in the scene. The results of the comparisons between the TDIP and Autoscope data reflect the importance of these factors. Table 2 gives the percentage of correspondence between the Autoscope and TDIP counts for each approach of the 12 unsignalized intersection cases that were studied. The percentage of correspondence varies widely from very poor levels (a low of 13 percent) to very good levels (90 percent and above). In general, those sites with good viewing angles (where good separation of vehicle images was achieved) and low pedestrian volumes, good correspondence levels (above 90 percent) were achieved. When viewing angles

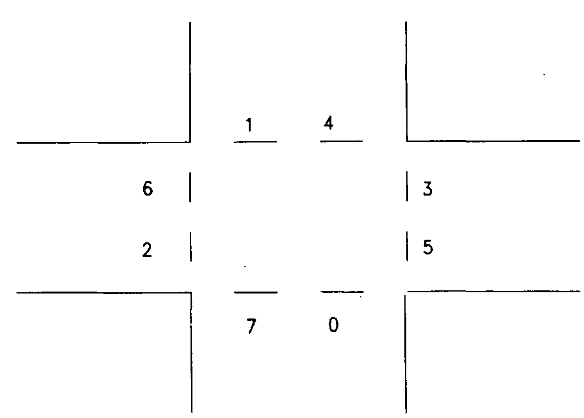

FIGURE 2 Sample detector layout for turning movement data (entry and exit detectors).

TABLE 2 Percentage Correspondence for Interrupted Flow Data Using TDIP and Autoscope

Site	Case Number	Percent Correspondence			
		Northbound	Southbound	Eastbound	Westbound
1	91031501	70	68	90	95
	91070801	75	53	92	95
	91071901	82	67	95	95
2	91031502	95	90	20	63
	91071101	93	73	40	84
	91071601	83	70	45	70
	91072301	92	63	20	64
3	91071201	86	13	80	56
	91072201	76	26	75	77
	91072501	71	35	89	80
4	91072601	-	53	94	94
	91073101	-	86	99	93

Note:
The data shown for each direction represent the percent correspondence between the TDIP data and the Autoscope data for the entire data collection period for each case cited. The TDIP data are used as the base for the correspondence calculation.

were poor or pedestrian volumes were high, the resulting correspondence between Autoscope and TDIP was poor.

Tables 3 and 4 present the turning movement volume counts from TDIP and from the Autoscope turning movement postprocessor algorithm for the unsignalized intersection sites for four 15-min periods. For the 21 cases in which the TDIP volumes exceeded 25 vehicles, the percentage deviations ranged from 0 to 21.6 percent, with a mean of 8.2 percent.

Discussion of Results

One of the major lessons learned during this part of the study was the importance of camera placement and viewing angle. To function successfully, the Autoscope system needs an unobstructed view of all vehicles passing through an intersection. In addition, detectors must be located such that the likely paths of pedestrians crossing through the intersection are avoided. If these criteria are satisfied, Autoscope can produce accurate approach volume and turning movement counts.

MEASUREMENT OF VEHICLE HEIGHT

Autoscope was not originally intended to measure vehicle size directly, but a process was developed as part of this research project to test how well Autoscope might perform this task. This information was of particular interest to the Idaho Transportation Department in monitoring the conformance of vehicles traveling through ports of entry in meeting legal size requirements.

TABLE 3 Turning Movement Count Comparisons, Eastbound

Case	Eastbound Left Turn		Eastbound Through Movement		Eastbound Right Turn	
	AutoTurn	TDIP	AutoTurn	TDIP	AutoTurn	TDIP
91071601	0	0	2	3	2	1
91071201	3	1	43	43	29	37
91070801	2	1	75	80	3	2
91071901	0	0	58	63	2	1
91072201	3	0	38	42	29	33
91072301	1	0	0	1	1	0
91031501	1	2	94	96	5	5
91031502	3	2	-	-	3	5
91072501	5	1	49	51	35	43
91072601	12	14	39	43	-	-
91073101	21	25	82	89	0	0

Note:
The data shown above are the 15-minute volumes collected directly by TDIP and produced by the turning movement post-processor using the Autoscope data (AutoTurn).

TABLE 4 Turning Movement Count Comparisons, Westbound

Case	Weastbound Left Turn		Westbound Through Movement		Westbound Right Turn	
	AutoTurn	TDIP	AutoTurn	TDIP	AutoTurn	TDIP
91071601	25	28	1	0	10	10
91071201	14	13	46	47	4	1
91070801	2	2	64	66	2	1
91071901	2	3	52	56	0	1
91072201	20	18	60	68	3	0
91072301	12	18	0	0	4	10
91031501	5	7	87	88	4	4
91031502	26	7	-	-	13	4
91072501	23	19	71	77	5	0
91072601	-	-	-	-	-	-
91073101	0	0	114	111	42	47

Note:
The data shown above are the 15-minute volumes collected directly by TDIP and produced by the turning movement post-processor using the Autoscope data (AutoTurn).

Objective of Measurement Test and Variable To Be Measured

The objective of this test is to determine how well Autoscope can measure vehicle height, and the variable to be measured is the height of the vehicle in meters.

Description of Measurement Methodology

The truck heights were measured manually by a standard measuring rod. Data were collected for 79 trucks, representing a variety of truck types.

Each truck was also videotaped as it passed through the port of entry. Reference points were marked on a utility pole at 0.3-m spacing in the camera field of view using colored tape. These points were used as a guide for locating video detectors for the Autoscope measurements. The detector placement and measurement method are shown in Figure 3. Autoscope recorded the time that each of these detectors were turned on and off by the passage of the truck image. The height was estimated by noting the maximum number of detectors that were activated during each truck passage. Effects of camera angle and site geometry were included in this measurement estimate.

Results

The times that each detector turned on and off were plotted for each truck. The result of this plot is a profile or representation of the visual image of the truck. A sample plot for one trucks is shown in Figure 4. The shape of the truck components, both the cab and the trailer, are clearly visible.

The field measurements were compared with the height measurements estimated by Autoscope. A histogram of the distribution of the measurement errors is given in Figure 5, which shows that for 70 of the 79 trucks studied, the measurement error between the actual measurement and the height measurement estimated from the Autoscope detector data is less than 10 percent.

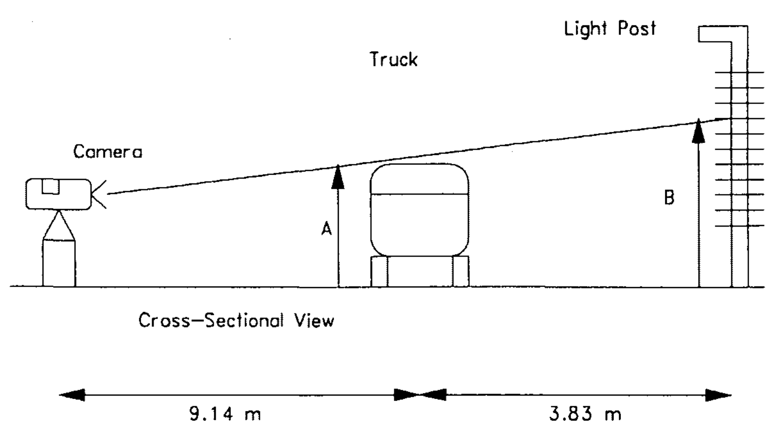

FIGURE 3 Truck height measurement method: actual truck height (A) is estimated from imagined height (B), using cross-sectional geometry shown.

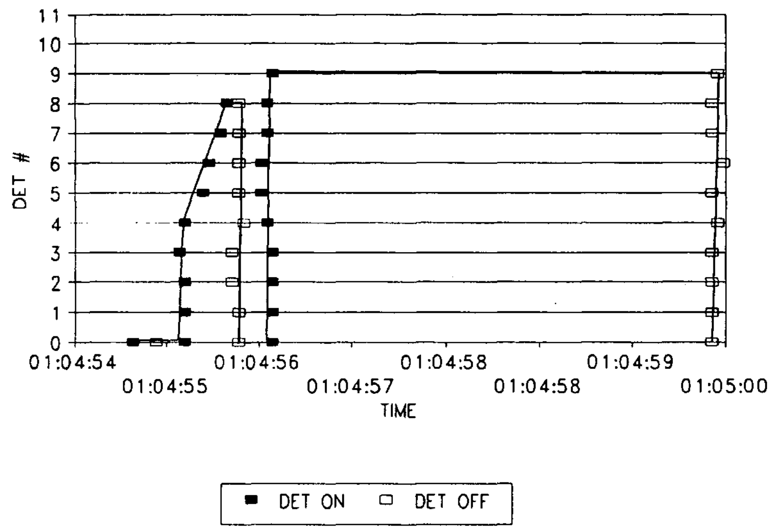

FIGURE 4 Sample truck profile, Truck 17.

Discussion of Results

An analysis of the measurements with the highest errors showed two common problems:

1. Autoscope detected the truck exhaust pipe, which was not included in the manual measurement, and thus overestimated the actual height.
2. Cylindrically shaped light-colored trucks (typically oil or agricultural tank trucks) were not properly detected by Autoscope, resulting in an underestimation of the height. The cause of this latter problem appears to be a some difficulty in differentiating the light intensity produced by the edge of the truck surface from the level produced by the pavement or background.

Figure 6 shows a plot of the actual measured height versus the height estimated by Autoscope, but it excludes the data for trucks that fit the two problem categories described earlier. The results shown in Figure 6 are very promising, particularly when it is noted that the detector grid consists of detectors that are spaced 0.3 m apart. Increased accuracy should be possible when more closely spaced grids are used.

MEASUREMENT OF VEHICLE LENGTH

Vehicle lengths were also measured using the Autoscope system. The results were compared with the data collected using standard manual field measurement methods.

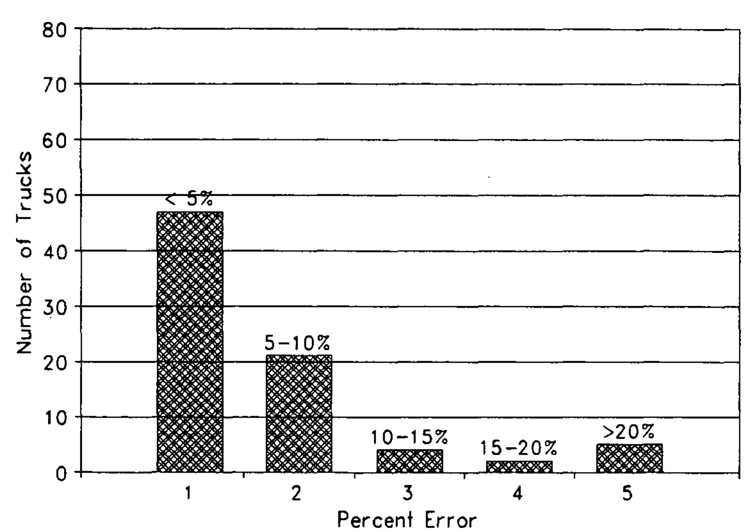

FIGURE 5 Distribution of percentage errors for truck height measurements.

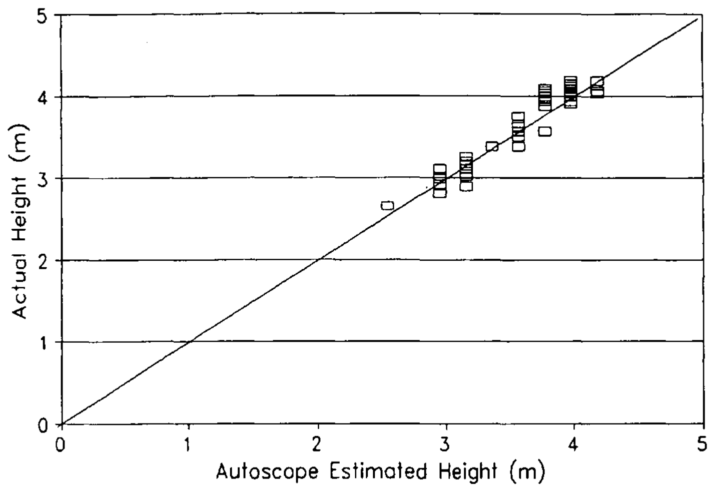

FIGURE 6 Comparison of actual and estimated truck heights, selected vehicles (shape and color) eliminated.

Objective of Measurement Test and Variable To Be Measured

The objective of this test is to determine the effectiveness of the Autoscope system in measuring vehicle length, and the variable to be measured is vehicle length in meters.

Description of Measurement Methodology

As in the vehicle height test, manual measurements of bumper-to-bumper vehicle length were made of 79 trucks at the Lewiston port of entry. A video camera was placed approximately 9.1 m from the truck lane to record the passage of the truck through the port of entry. The camera angle was established so that the entire truck image could be captured. The actual view of the truck passage was therefore not straight on, but somewhat oblique.

To collect the Autoscope data, vertical video detectors were plotted at 1.5-m intervals along the path of the truck. As the truck progressed through the field of view, the detectors were activated in order from the left to the right of the field. The detector activation data are plotted on a time-distance diagram to indicate the progression of each truck through the detector chain. An example of this procedure is given in Figure 7. The horizontal line noted by "Estimated Truck Length" in the figure represents the maximum number of detectors simultaneously activated during the time of truck passage.

Results

The maximum number of detectors activated by a truck passage was plotted against the field measured length for each truck. This plot is shown in Figure 8. The correlation coefficient between the actual length of the truck and the number

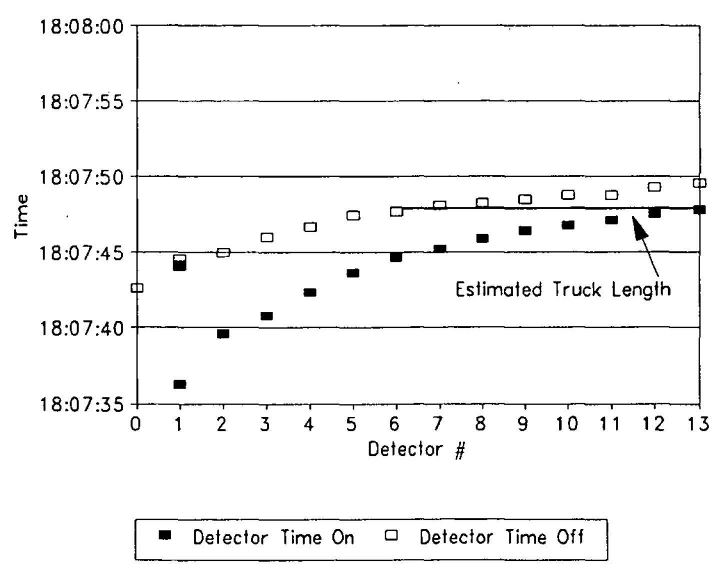

FIGURE 7 Truck length measurement.

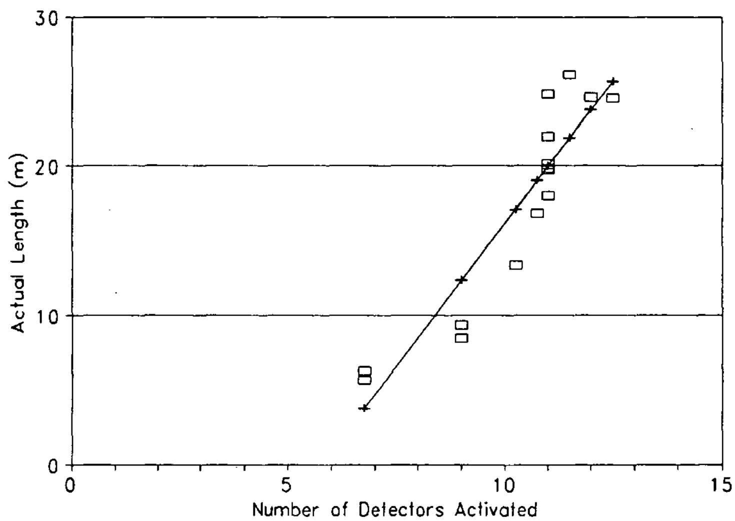

FIGURE 8 Measured truck length versus number of activated detectors.

of simultaneously activated detectors is 0.85, indicating a good level of correspondence between the actual length measurement and the Autoscope estimate. This is particularly encouraging since the detector grid (with 1.5 m between each detector) was relatively coarse. Unfortunately, a direct estimation of the length using the Autoscope data was not possible because of the obliqueness of the camera angle.

Discussion of Results

The test shows that there is potential for using video imaging to measure truck length. However, better camera placement will be required to obtain a more direct view of the truck image passing through the field of view. An improved orientation of the truck image will allow a calculation of the length from the site geometry.

SUMMARY AND CONCLUSIONS

The purpose of the paper is to present the results of a series of tests of the Autoscope 2002 video imaging system in collecting transportation data. The following is a summary of the primary results of this study:

- Proper camera placement and the establishment of an obstructed view of all vehicles passing through the field of view are both essential to the successful collection of data using video imaging techniques. Occluded views of vehicles as well as the presence of extraneous images (such as pedestrians) may degrade the quality of the results obtained.
- Postprocessors have the potential to increase significantly the value of the data produced by a video imaging system by providing the analyst with additional information about a transportation system.
- The results of the four measurement tests are promising:
  - –In the measurement of freeway traffic volume counts, Autoscope produced data that varied from 0.6 to 9.3 percent of manually collected data, when proper camera angle and detector configurations were selected.
  - –Autoscope produced approach flow rates with correspondence levels of 90 percent or greater with manually collected data for unsignalized intersections when suitable camera angles were selected and when pedestrian volumes were low.
  - –The turning movement postprocessor produced volume estimates that were within 8 percent of the manual counts.
  - –Height estimates generated by Autoscope were within 10 percent of the manual height measurements for 70 of the 79 trucks studied.
  - –The correspondence between the manual length measurement and the number of simultaneously activated video detectors showed a correlation coefficient of 0.85.

Clearly more work needs to be done in the validation of this system in a wider variety of situations and in the development of suitable postprocessors so that additional transportation data can be collected more efficiently and more accurately. But these initial results show good potential for machine vision as a tool in the collection of transportation data.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Idaho Transportation Department and the University of Idaho during the research project described in this paper.

REFERENCES

1. Michalopoulos, P., B. Wolf, and R. Benke. Testing and Field Implementation of the Minnesota Video Detection System. In *Transportation Research Record 1287*, TRB, National Research Council, Washington, D.C., 1990.
2. Hockaday, S. *Evaluation of Image Processing Technology for Applications on Highway Operations.* Contract 51K287, TR 91-2. California Polytechnic State University, San Luis Obispo, 1991.

3. Taff, S., and W. L. Garrison. *Image Processing and Highway Operations*. Institute of Transportation Studies, University of California, Berkeley, 1990.
4. Gilbert, R. K., and Q. Holmes. Dynamic Traffic Information from Remote Video Monitors. *Proc., Vehicle Navigation and Information Systems Conference,* IEEE, Dearborn, Mich., Oct. 1991.
5. Vieren, C., J. P. Deparis, P. Bonnet, and J. G. Postaire. Dynamic Scene Modeling for Automatic Data Extraction. *Journal of Transportation Engineering*, Vol. 117, No. 1, 1991.
6. Sethi, I. K., and W. L. Brillhart. *Proc., Vehicle Navigation and Information Systems Conference,* IEEE, Dearborn, Mich., Oct. 1991.
7. Troutbeck, R. J., and J. S. Dods. Collecting Traffic Data Using the ARRB Video Analysis Data Acquisition System. *Proc., 13th ARRB/5th REAAA*, 1986.
8. Taylor, M. A. P., W. Young, and R. G. Thompson. Headway and Speed Data Aquisition Using Video. In *Transportation Research Record 1225*, TRB, National Research Council, Washington, D.C., 1989.
9. Michalopoulos, G. *Automatic Incident Detection Through Video Image Processing*. Department of Civil Engineering. University of Minnesota, Minneapolis/St. Paul, 1991.
10. *NCHRP Research Results Digest 181: Video Image Processing for Evaluating Pavement Surface Distress*. TRB, National Research Council, Washington, D.C., Dec. 1991.
11. Koutsopoulos, H. N., and I. E. Sanhouri. Methods and Algorithms for Automated Analysis of Pavement Images. In *Transportation Research Record 1311*, TRB, National Research Council, Washington, D.C., 1991.
12. Lu, Y. J., X. Yuan, and C. D. Yan. Application of Image Sequence Analysis to Pedestrian Volume Measurement at Intersections. Presented at 71st Annual Meeting of the Transportation Research Board, Washington, D.C., 1992.
13. Boesen, A., M. Kyte, and B. Rindlesbacher. *Traffic Data Input Program: Users Manual and Software*, Version 3.0. Department of Civil Engineering, University of Idaho, Moscow, March 1991.

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

Computer-Administered Surveys at Honolulu International Airport

C. S. PAPACOSTAS, NICOLAOS E. SYNODINOS, AND GLENN M. OKIMOTO

The management of the Honolulu International Airport (HIA) implemented a computerized user opinion survey system employing a touch-screen interface to elicit, on a continuing basis, data relating to user satisfaction. The substitution of the new system for paper-and-pencil surveys raised several questions, including whether the paper-and-pencil questionnaires and the touch-screen system obtained responses from the same subject pool, whether the respondents gave comparable answers to the same questions posed by way of the two survey modes, and whether any issues associated with the new method needed special attention. A study addressed these questions through (a) an exposure count to determine the number of persons exposed to the system and their characteristics, (b) a structured observation study to associate the specific categories of respondents with the response patterns recorded by the touch-screen system, and (c) an investigation of the effects of survey administration method and sample selection method. On the basis of the observation study, rules were established for filtering out of the computer-generated files the records created by unreliable self-selected respondents to the touch-screen survey. The study found that the sample selection protocol had a significant effect on the survey responses. Whereas the specific recommendations offered are applicable to HIA, the implications of the study will be useful to those contemplating similar uses of the new technology.

The degree to which users are satisfied with the services offered at major facilities such as shopping centers and airports is important to the management of these facilities. Collected on a continuing basis, related information can help management identify problem areas and assess the success or failure of operational decisions. Until recently, this task was accomplished through the conduct of user surveys that were of the paper-and-pencil variety. These surveys would typically be repeated at regular or irregular intervals to capture changes in user perceptions over time.

Computer-administered telephone surveys relying on mainframe systems can be traced to the early 1970s (*1–3*). With the recent proliferation of microcomputer hardware and software, computer-aided interviewing techniques are becoming increasingly accessible and cost-effective. In particular, touch-screen systems, where the respondent answers by pressing against active regions of the computer monitor, have a good potential to reduce the cost and inconvenience of paper-and-pencil methods.

C. S. Papacostas, Department of Civil Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, Hawaii 96822. N. E. Synodinos, Department of Marketing, University of Hawaii at Manoa, 2404 Maile Way, Honolulu, Hawaii 96822. G. M. Okimoto, Department of Transportation, State of Hawaii, 869 Punchbowl Street, Honolulu, Hawaii 96813.

Typical applications of touch-screen systems at major public facilities entail the presentation of information to the user, such as the provision of directory assistance at shopping centers. A touch-screen survey method has been implemented at the Honolulu International Airport (HIA) to collect data on user satisfaction. As implemented at HIA, the system relies on self-selected participants who complete the survey unsupervised. Because of the novelty of the application, there is a need to understand how respondents interact with the new system and to assess its comparability to the earlier paper-and-pencil version used at HIA. To do this, two observational studies (an exposure count and a structured observation study) were conducted. The effect of two sample selection methods (self-selection versus facilitator-selection) and the effect of the two methods of survey administration (paper-and-pencil versus computer touch screen) were also compared. The findings of the study described in this paper will be useful to those contemplating similar real-world applications of this technology.

STUDY CONTEXT

One of the busiest airports in the United States, HIA is the major hub of the statewide airport system operated by the Hawaii Department of Transportation (HDOT). In a typical year, HIA processes more than 20 million enplaning and deplaning passengers (*4*).

Being responsive to the needs of airport users, the Programming, Planning, and Budgeting Office and the Airports Division of HDOT have been conducting user opinion surveys employing paper-and-pencil instruments administered to facilitator-selected respondents. In 1987 they decided to explore the possibility of using a computerized interviewing system. They obtained a touch-sensitive computer screen connected to a stand-alone microcomputer and hired an independent consultant to develop an illustrative survey to test the feasibility of this approach. On the basis of preliminary comparisons (*5*), HDOT expressed some concerns about the equivalence and compatibility of the data collected via the paper-and-pencil version and the results obtained via the touch-screen system. A research project was awarded to the University of Hawaii at Manoa to address the questions of whether (a) the previous paper-and-pencil questionnaires and the touch-screen system obtained responses from the same subject pool, (b) the respondents gave comparable answers to the same questions posed by way of the two survey modes, and (c) any issues intrinsic to the new method needed special attention.

The overall study was divided into three parts: an exposure count was conducted to determine the number of persons passing by the system and their characteristics, the number of computer records created during the same time, and the ways in which respondents interacted with the system. A structured observation study associated specific categories of respondents with the response patterns recorded by the touch screen–based system. Of particular relevance was the establishment of criteria for filtering out of the data files the records created by unreliable respondents. These included children who appeared to be using the system as a substitute for an arcade game. Finally, the findings of the structured observation study were incorporated in the selection of three respondent samples to investigate the effects of survey mode and sample selection method.

PAPER-AND-PENCIL INSTRUMENT

The paper-and-pencil questionnaire used for the study was a slightly modified version of the instrument developed by HDOT. Minor changes in the wording and sequence of the questions were made to render it identical to the computer version that was implemented at the airport. The three-page questionnaire consists of 18 questions, some of which are designed to be skipped by certain respondents.

The opening question asks why the respondent is at the airport; the choices of response are that they are leaving, arriving, picking someone up, or seeing someone off. Depending on the answer, the respondent is directed to the appropriate branches of the questionnaire. Departing passengers, for example, are asked to specify their destinations and to rate their experience at the agricultural and security checkpoints that apply to their situations. By contrast, arriving passengers are asked about their points of origin but not about the elements not experienced on that day.

All respondents are then asked to rate 13 aspects of the airport, including their overall opinion of HIA on 11-point scales ranging from 0 (very poor) to 10 (very good) and including a response of "don't know." The questionnaire then elicits several demographic characteristics, including the respondent's education level, age, household income, and place of residence. The respondents are then asked to specify the frequency with which they visited HIA in the last year and whether they had answered the survey in the preceding 3 months. Finally, an open-ended question asks for comments.

TOUCH-SCREEN INSTRUMENT

The touch-screen questionnaire is structurally identical to the paper-and-pencil survey. However, it incorporates colorful graphics, animation, and an alternative method of entering responses. For example, to answer questions relating to geography (i.e., origins and destinations), a respondent merely presses directly the desired location on a displayed color-coded map of the world rather than picking from a printed list of geographical regions. Another difference is that the computer program automatically skips to the appropriate branch given a particular response, preventing respondent errors in this respect. Comments are elicited via a simulated keyboard that is displayed on the computer screen. This requires the respondent to press individual simulated keys to form a response, as opposed to the more familiar written response required by the paper-and-pencil version. Discovering whether the collective differences between the two systems affect the obtained data was one of the objectives of the study.

The computer version incorporates two time-outs that operate as follows: when a duration equal to the first time-out elapses before the respondent answers a question, an audio warning is emitted and a prompt screen encourages the respondent to continue. If no response is given during an additional interval equal to the second time-out, the partly completed record is written to file and the system is returned to the opening screen.

The data files created by the system follow a basic structure. A record containing the date and the time is written when the computer system is started. This record is followed by normal transaction records, each consisting of 183 characters, beginning with its starting time, and ending with its completion time. It includes the responses to the 18 questions that make up the questionnaire and allows for a 120-character comment field. Aborted records can be identified by the presence of the code "8" in the remaining response fields.

EXPOSURE COUNT

An observer was positioned at some distance from the location of the computer system and recorded the volume and direction of people passing by the system. The observed persons were categorized as adult man, adult woman, male airport employee, female airport employee, and child. The category "undetermined" was used at certain times when the observer was unable to categorize individual members of very large groups. The beginning and end of each exposure count session were marked in the computer data file by creating easily identifiable opening and closing records, allowing an exact and objective count of the number of computer records created during the observation periods. This number was contrasted with the number of respondents, which as explained later, was not identical to the number of records created.

Eight sessions spread over 3 days were held, totaling 15 hr and 27 min. During this time, 9,864 persons passed directly in front of the system and 176 computer records were created. There were approximately equal numbers of adult men and women in the observed sample. Children composed fewer than 10 percent of the persons observed but they tended to visit the system more than once.

Only a small fraction of the observed persons interacted with the touch-screen system. Overall, an average of 1.78 records were created per 100 persons exposed to the system. Considering the high volume of people visiting HIA, however, the system produces an enormous amount of data that need processing. Approximately a third of the records created during the exposure count were started but not completed to any degree.

During the exposure count, the observers noticed that the interaction of the respondents with the machine and the creation of individual records was a complex matter that deserved further investigation to ensure the proper analysis and interpretation of the data obtained in this manner. For example, some questionnaires were begun by one respondent but were completed by another. The reliability of the raw data collected via the touch-screen system could, therefore, be compromised. To address this problem, a detailed association between the persons contributing to the creation and to the content of these records was warranted, which led to the structured observation study and the record filtering rules described next.

STRUCTURED OBSERVATION

Method

The observational study was conducted on 10 days and yielded information corresponding to 224 records. An observer was stationed at a location with a clear view of the computer screen that also ensured unobtrusiveness. The observer recorded the characteristics of each respondent on a one-page special form. One form was sufficient in cases in which a respondent was singly responsible for the creation of a single transaction record in the data file. On the other hand, several forms were needed to keep track of multiple respondents to the same transaction record. In the cases of such "chained" records, the question posed on the computer monitor at the time when each respondent took control of the system was noted in the corresponding form. The beginning and ending times of each transaction were also noted. This permitted a one-to-one matching between the transaction records produced by the touch-screen system and the associated information obtained by the observer. Included among the respondent characteristics were gender and estimated age. The observed behaviors about the pattern of interaction of each respondent with the touch-screen system were also entered on the form.

As explained earlier, the number of respondents who participated in producing these records was different from the number of created records. In addition to chained records, some respondents completed the survey more than once ("repeats"). On this basis alone, it would be erroneous for an analyst of the data to assume a one-to-one association between a record collected via the touch-screen survey and an individual respondent. A record filtering process that can minimize the proportion of inappropriate records before data analysis is desirable. Given the potential unreliability of records generated by children, the use of age-related criteria for filtering was investigated.

Record Filtering

Only six age categories were included in the tally sheets completed by the observer rather than the eight categories found in the data file records because the estimation of age by observation involves judgment. The observed age brackets included the ranges 30 through 49 and 50 through 69 that, in the computer-written records, consist of two categories each.

The 10–19 age category included many persons between 10 and their early teens. Through annotation, the observer tried to capture the size of this subgroup and found that about 50 percent of the respondents observed to fall in the 10–19 category were estimated to be 16 or under. The presence of these respondents in the 10–19 category could possibly render the raw aggregated responses by the entire group unreliable.

Table 1 cross tabulates the ages of the respondents as estimated by the observer (rows) against the ages reported by the respondents in the data file records (columns). Two additional columns are included, one for transactions that were aborted before any of the survey questions were answered and one for partially completed records. Each cell of this table contains three entries corresponding to individual respondent records, chained records, and repeated records.

The region of Table 1 around the main diagonal that is delineated by heavy lines represents the region of "reasonable match" between the observed and recorded ages of the re-

TABLE 1 Observed Versus Reported Age

OBSERVED AGE	REPORTED AGE < 10	10 - 19	20 - 29	30 -39	40 - 49	50 - 59	60 - 69	> 70	ABORTED RECORDS	PARTIAL RECORDS
< 10	2[a] 2[b] 5[c]	3 1 3			1 0 0			3 0 2	20 0 6	5 3 0
10 - 19		32 2 1	7 0 1	2 0 4	3 0 0			2 1 0	10 4 5	4 2 1
20 - 29		1 0 0	19 3 0	8 4 0				1 0 0	6 0 0	3 2 0
30 - 49	2 0 0	1 0 0	1 0 0	6 1 0	9 2 1	2 2 0			5 0 0	3 1 0
50 - 69						2 0 0			3 0 0	
≥ 70							1 0 0			

[a] Records created by individual respondents
[b] Chained records
[c] Repeated records

spondents. It shows that respondents are generally truthful in reporting their ages. Interestingly, all of the observed records that lie outside and to the left of the solid outline were cases in which older adults directed children to enter suggested responses. Some of the outliers to the right of the solid outline were also of this "instructional" type, but apparently in these cases the older adults directed the children to enter the adults' ages. The most pronounced age discrepancy is seen in the record category of 70 years of age and older, in which all observed entries were made by children under 10 and, particularly, by those on the younger side of the 10–19 group. Children in these brackets showed a tendency to repeat the survey and to also produce aborted and partial records at a higher frequency than adults.

Concern about the reliability of the responses given by children had motivated the HDOT to eliminate the under-10 category from their initial analyses of the data obtained via the touch-screen system (*5*). The findings of the observational study reinforce the need for this practice. Moreover, it appears prudent to also exclude the 10–19 and the over-69 age brackets and all partially completed records. This practice will neither ensure the reliability of all retained records nor guarantee the removal of only unreliable records, but it will eliminate a good portion of the unreliable records. With data filtering, the presence of children among the users of the survey is not as critical as it would first appear because they tend to abandon the survey midstream. Their transactions would thus be removed by retaining only completed surveys for subsequent analyses. The same is true with many repeated and chained records.

The effect of the filtering rule based on reported age on the computed mean ratings given to the 13 airport attributes was examined by a series of *t*-tests computed using sample sizes ranging from 2,500 to 3,400 responses to the rating scales. These analyses showed that the average ratings of the airport when using unfiltered data are different from the ratings obtained when filtering the raw data set. Consequently, the conclusions drawn regarding the airport would be different depending on which of the two data sets is used. All available evidence points to the reasonableness of using the filtered data set.

Other Filtering Options

Other ways of identifying potentially unreliable records were also investigated. These included attempts to discern internal inconsistencies in the responses, specific response patterns, and parsing of the comment field. Several special processing routines using the PROLOG language were written for this purpose. The general conclusion was that the vast majority of apparently inappropriate records would be removed by the age filtering rule just discussed. Nevertheless, removal of records containing indecorous or infantile comments and records indicating that the respondent has answered the questionnaire previously can eliminate additional unreliable records.

The application of these filtering criteria would define a "reference group," the responses of which could be used to track changes in airport user satisfaction over time. It can also capture the reactions of users to management actions intended to improve the airport.

METHOD AND SAMPLE EFFECTS

Purpose and Scope

The last part of the study examined the degree to which the method of survey administration influences the obtained responses. From a practical point of view, the comparison of interest to the airport's management entailed the paper-and-pencil survey versus the touch-screen method. However, as implemented at HIA, the former involved a facilitator-selected sample of respondents, whereas the latter allowed respondents to be self-selected. To isolate the effect of sample selection from the effects of survey administration method, three combinations of sampling and survey administration were contrasted: the computer-administered questionnaire completed by a randomly selected group (CA/R), the computer-administered questionnaire completed by a self-selected group (CA/S), and the paper-and-pencil questionnaire completed by a randomly selected group (PP/R).

This allowed for three sets of comparisons: the comparison of CA/R and CA/S would shed light on the effect of the sample selection method on the results obtained by the touch-screen system; the comparison of CA/R and PP/R would permit the examination of the compatibility of the two methods of survey administration, holding the sampling method constant; and the comparison of CA/S and PP/R would provide information on the results when both the administration and sampling methods varied.

Sample Selection Procedures

A systematic procedure was followed in selecting the three samples of respondents to ensure that the location and time frames were similar for the three cases. The two random samples were chosen by survey facilitators. The third adult passing by an unoccupied facilitator was asked to participate and, if agreeable, was either given a clipboard with the paper-and-pencil questionnaire or directed to the computer. The survey facilitators kept a contact record sheet for each person approached. Entered on the contact record sheet were the survey method, the date and time, the subject's gender and estimated age, and whether the person agreed or refused to participate. Numbering of the contact records facilitated the matching of particular transactions to individual respondents.

Thirty-six data collection sessions were conducted over 79 days to yield a sample size of 642 respondents: 330 completed paper-and-pencil questionnaires, and 312 completed valid computer records. The data collection sessions covered the hours from 8:00 a.m. to midnight, coinciding with the bulk of airline operations at HIA. χ^2 tests showed that the persons approached to do the survey by the two methods were statistically similar in terms of age, gender, and refusal rates. The age and gender similarity extended to the two groups that agreed to participate. Thus the sample selection process was successful in creating equivalent random samples (CA/R and PP/R) in terms of these variables.

The third sample that completed the touch-screen survey unsupervised (CA/S) was drawn from the offloaded computer data files. The CA/R record sets were taken from approximately the same time frames as the facilitator-selected sets.

On the basis of the results of the earlier observational study, the CA/S sample was reduced by filtering out incomplete records and records indicating age categories of younger than 20 and older than 69. For the sake of consistency, the same filtering was applied to the two randomly selected samples (CA/R and PP/R) as well.

Analysis and Findings

The responses obtained by the three groups were compared pairwise with respect to their responses to each of the 13 facility rating scales, the respondents' characteristics, and their propensity to offer comments. The comparison of rating responses was done via a series of *t*-tests, whereas the comparisons of respondent characteristics and their propensity to comment were done using χ^2 tests. Following is a summary of the principal findings of these comparisons (*6*). In the following sections, significance refers to the 0.05 level.

In terms of the available demographic characteristics, the groups associated with the three methods (i.e., CA/R, CA/S, and PP/R) were found to be of similar reported incomes and levels of educational attainment. The two randomly selected groups (CA/R and PP/R) were also similar in terms of age. Both included relatively older respondents as compared with the sample of self-selected respondents. The reported ages of the two randomly selected groups were in agreement with the ages estimated by the survey facilitators. This reinforces the finding of the earlier observation study, which found the reference group of respondents to be truthful in this respect.

The comparisons between the two randomly selected groups with respect to their responses to the 13 rating scales showed the PP/R slightly higher than the CA/R. However, statistical significance was reached only in 2 of the 13 airport attributes (Table 2). The preponderance of the findings supported the general conclusion that the responses obtained via the two instruments (i.e., computer and paper and pencil) were not different when respondents were selected randomly.

The self-selected group that answered the touch-screen survey, on the other hand, was found to be different than the two randomly selected groups. Table 2 indicates that the CA/S group gave significantly lower ratings than the PP/R group on 9 of the 13 airport attributes and significantly lower ratings than the CA/R group on 5 of the 13 scales.

TABLE 2 Comparability of Survey Methods

SCALE	CA/R vs. PP/R	CA/S vs. CA/R	CA/S vs. PP/R
Appearance	n.s.	n.s.	n.s.
Airport Roads	n.s.	n.s.	CA/S < PP/R
Airport Parking	n.s.	n.s.	n.s.
Shuttle Bus	n.s.	n.s.	n.s.
Loading Zones	n.s.	n.s.	CA/S < PP/R
Directional Signs	n.s.	CA/S < CA/R	CA/S < PP/R
Baggage Areas	n.s.	n.s.	CA/S < PP/R
Restaurant/Snackbars	CA/R < PP/R	n.s.	CA/S < PP/R
Gift Shops	CA/R < PP/R	CA/S < CA/R	CA/S < PP/R
Public Conveniences	n.s.	CA/S < CA/R	CA/S < PP/R
Visitor Information	n.s.	CA/S < CA/R	CA/S < PP/R
Cleanliness	n.s.	n.s.	n.s.
Overall	n.s.	CA/S < CA/R	CA/S < PP/R

n.s. = not significant at the 0.05 level.

With respect to the propensity of respondents to offer comments, the study found a great variability between the three methods: comments were entered by 50 percent of the self-selected respondents who used the touch-screen system, by 37 percent of the paper-and-pencil survey respondents, and by only 15 percent of the randomly selected respondents who answered the computer survey. Comments written on the touch-screen system were more difficult to decipher than those obtained via the paper-and-pencil questionnaire, partly because of difficulties in using a simulated keyboard on the computer monitor. However, the high volume of comments obtained under normal operation of the touch-screen system (CA/S) makes it a source of useful and continuously collected information, assuming that obviously improper comments are judiciously filtered out. In fact, the "suggestion box" feature of the touch-screen system is one of its strong points.

DISCUSSION OF RESULTS AND CONCLUSIONS

The use of computer-interactive survey systems in general and systems employing a touch-screen interface in particular is expected to increase considerably. For major activity centers such as airports, these systems offer a relatively inexpensive means of collecting, on a continuing basis, data relating to user perceptions and opinions.

Besides providing timely results, properly programmed computerized questionnaires can minimize certain respondent errors. In particular, providing automatic branching to the appropriate sections of the questionnaire on the basis of specific responses eliminates the difficulties encountered by respondents to paper-and-pencil questionnaires in this respect. This advantage of computer interviewing becomes more pronounced as the complexity of the questionnaire increases. In fact, since the skips are transparent to the respondent, it is possible to construct more refined questionnaires without fear that the complexity of the questionnaire will overwhelm the respondents. Furthermore, questions in computer interactive surveys "can be grouped either in a more intuitive manner or per the dictates of the research design" (*7*).

Another advantage of computerized questionnaires is that they can provide data in the form of ASCII files that are ready for further processing, which eliminates the step of transcribing the survey responses into computer files. Whether the translation is done by keypunching or by scanning of machine-readable forms, this operation often involves errors requiring costly and time-consuming checking.

The use of a touch screen–based interviewing system offers another advantage relevant to surveys related to the spatial aspects of transportation systems. Geocoding is greatly enhanced through the use of an on-screen touch-sensitive map. For the study described in this paper, the number of origins and destinations of respondents consisted of seven regions of the world (e.g., U.S. mainland, Canada, other Hawaiian islands, etc.). In the paper-and-pencil version, the respondents were asked to pick from a list, whereas touch-screen system respondents were simply required to touch the corresponding geographical area on a color-coded map. The limitation on the number of response choices to seven was imposed by a need to keep the paper-and-pencil questionnaire to a reasonable length. A more elaborate implementation would permit a much finer designation of geographical zones. That case is typically treated by paper-and-pencil instruments as an open-ended question that requires extensive and costly geocoding.

The quality and reliability of the raw data obtained via the touch-screen system would depend, among other factors, on the characteristics of the application and on the respondent selection procedures. For the HIA application described in this paper, a self-selection protocol was deemed appropriate by the airport's management because of a desire to collect data continuously while avoiding the costs associated with supervised surveys. The reliability of the computer-created records, however, could be judged only minimally. Consequently, HIA management wanted to assess the potential effect of unreliable respondents on the results obtained via the touch-screen system and possible ways to reduce this effect without requiring the constant attention of a survey supervisor. Also of concern was the methodological question of the comparability of the results derived through the new system to those obtained via commonly administered paper-and-pencil surveys that rely on facilitator-selected random samples.

To address the first issue, a structured observation study associated information about a sample of self-selected respondents to the computer records they created. On the basis of the comparison between the two, several filtering rules were established that can remove a considerable number of spurious records. In order of importance, these rules were to eliminate records reporting ages under 20 and over 69, to retain only complete records, and to eliminate records containing incongruous and other improper comments.

The issue of comparability between paper-and-pencil surveys relying on facilitator-selected random samples and touch screen–based surveys of self-selected respondents required isolation of the effects of survey administration method (paper-and-pencil versus computer) and the effects of sample selection (random versus self-selection).

The analysis showed that for similar facilitator-selected random samples the responses obtained via the touch-screen system were statistically comparable with the responses derived from the structurally identical paper-and-pencil survey. The introduction of a self-selected group of respondents, however, resulted in statistically significant differences. The largest differences were found between CA/S and PP/R, the two methods that are customarily employed at the HIA. An implication of practical relevance to the management of the HIA is that the two commonly used methods are not directly comparable. Proper comparisons of the PP/R ratings obtained in earlier years with the CA/S obtained after the installation of the touch-screen system would require adjustments of the mean ratings of one or the other. This would avoid, for example, an erroneous conclusion that the level of satisfaction of airport users declined in 1987, the year when the touch-screen system was put into operation. The ratios of the mean ratings obtained in the study reported here can serve as first-cut adjustment factors.

Although significant differences were found between the two facilitator-selected samples on one hand and the self-selected sample on the other, there are advantages to continued operation of the touch-screen system on a self-selection basis. As operated at the HIA, the system can record the perceptions of airport users continuously, 24 hr a day, and in real time. It is possible, therefore, to assess the reactions of airport users to particular events or changes in airport operations and conditions as they occur. For example, overbooking by airlines during certain periods, uncomfortable conditions during parts of the year, and responses to schedule changes can be detected through the contemporaneous data obtained by the system. Some of these occurrences would be unanticipated and very difficult to capture by occasional supervised surveys that require advanced planning.

In addition, some planned one-time surveys could happen to coincide with the occurrence of atypical conditions at the airport and thus could result in conclusions that may not be applicable under normal circumstances. Nevertheless, the use of the touch-screen system with self-selected respondents does not obviate occasional facilitator-supervised surveys, whether administered by computer or not. Such surveys would be necessary to address issues not covered by the touch-screen system and also in control studies similar to that described in this paper.

Whereas these findings are directly applicable and relevant to the specific system installed at HIA, a transferable lesson derived from this study is that the raw data obtained by similar computer-administered surveys should not be accepted at face value. Instead, they should be explicitly evaluated within their specific contexts. Collectively, the accumulation of application-specific experiences will enhance the understanding and effective use of emerging survey research technologies.

ACKNOWLEDGMENTS

The material presented here is based on research funded by HDOT. The authors would like to thank Chuck Fisher of Magnus Communications Design, Inc., and Myra Tamanaha of HDOT.

REFERENCES

1. Nicholls, W. L., and R. M. Groves. The Status of Computer Assisted Telephone Interviewing: Part I: Introduction and Impact on Cost and Timeliness of Survey Data. *Journal of Official Statistics*, Vol. 2, 1986, pp. 93–115.
2. Fink, J. C. CATI's First Decade: The Chilton Experience. *Sociological Methods and Research*, Vol. 12, 1983, pp. 153–168.
3. Shanks, J. M. The Current Status of Computer-Assisted Telephone Interviewing: Recent Progress and Future Prospects. *Sociological Methods and Research*, Vol. 12, 1983, pp. 119–142.
4. *The State of Hawaii Data Book, A Statistical Abstract.* Hawaii Department of Business, Economic Development, and Tourism, Honolulu, 1991.
5. Okimoto, G., M. Tamanaha, and E. Asato. *Measuring the Effectiveness of Hawaii's Airports.* Hawaii Department of Transportation, Honolulu, 1987.
6. Synodinos, N. E., C. S. Papacostas, and G. M. Okimoto. *Evaluation of Computer Interactive Data Collection at the Honolulu International Airport.* Contract 23181. University of Hawaii at Manoa, Honolulu; Hawaii Department of Transportation, Honolulu, 1989.
7. Synodinos, N. E., and J. M. Brennan. Computer Interactive Interviewing in Survey Research. *Psychology and Marketing,* Vol. 5, 1988, pp. 117–137.

Opinions expressed herein are the authors' and do not reflect an official position of either the Hawaii Department of Transportation or the University of Hawaii at Manoa.

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

Traffic Data Collection Using a Computerized Data Acquisition System

BRIAN A. MOEN, JAMES A. KOLLBAUM, JAMES A. BONNESON, AND PATRICK T. MCCOY

Traffic characteristics such as headway, flow rate, speed, acceleration, wheelbase, and lane position have traditionally been collected using manual procedures, which are labor- and time-intensive and generally allow only one characteristic at a time to be collected. However, a computerized data acquisition system (CDAS) using tape switch sensors on the roadway will allow many traffic characteristics to be collected simultaneously. CDAS has three major advantages over a manual collection system: it requires fewer total person hours to collect the data, it offers improved measurement precision and accuracy, and it collects data continuously. An overview of a CDAS is presented, the methods used to extract the desired traffic characteristics from CDAS data are discussed, and insights and experiences obtained using a CDAS are offered.

Traffic characteristics such as headway, flow rate, speed, acceleration, wheelbase, and lane position have traditionally been collected using manual procedures. In this context, a manual procedure would include any study procedure wherein the data are collected in the field or extracted from videotape or film in a visual manner. Most of these procedures are labor- and time-intensive and generally allow only one characteristic at a time to be collected. However, a computerized data acquisition system (CDAS) using tape switch sensors on the roadway will allow many traffic characteristics to be collected simultaneously.

A CDAS has three major advantages over a manual collection system: (a) it requires fewer total person hours to collect the data, (b) it offers improved measurement precision and accuracy, and (c) it collects data continuously. Time requirements (in terms of total person hours) for data collection can vary significantly with the type of traffic characteristic being studied and the type of study procedure. For example, a manual study of driver headways will typically require two persons for each lane studied. In contrast, a CDAS-based study would require a study team of four to six persons to set up the study and one person to monitor the system after it has been set up, regardless of the number of lanes studied.

Time requirements for data reduction can also vary significantly with the type of study. In general, a CDAS automatically creates a computer file of the raw data, whereas with manual methods the data are recorded on field data sheets for later manual entry into a data base. Data from manual methods, however, generally are in the proper format for statistical analysis, whereas the CDAS data are often in a very basic form that requires some data reduction before any analysis.

Civil Engineering Department, University of Nebraska–Lincoln, Neb. 68588.

Measurement precision and accuracy are much better with a CDAS than with manual methods. Improved precision stems from the ability of the CDAS to measure event times to the thousandth of a second. Manual traffic studies that measure timed events have lower precision because they are limited by the reaction time of the human observer. Manual methods also have less accuracy because of the human inability to maintain an alert state during long-term field studies. As a result of these deficiencies, manual methods cannot accurately measure complex traffic characteristics such as acceleration and lane position.

This paper presents an overview of a CDAS, a discussion of the methods used to extract the desired traffic characteristics from CDAS data, and some insights and experiences obtained by the authors in using a CDAS. The CDAS system described in this paper has been used successfully by the authors on several projects since it was assembled 2 years ago. These projects have been directed toward measuring traffic characteristics at signalized intersections. As a result, much of the discussion is directed toward CDAS applications to intersections. However, it should be noted that most of the concepts in this paper can easily be extended to other locations (e.g, midblock arterial segments, tangent highway sections).

SYSTEM CHARACTERISTICS

The CDAS used by the University of Nebraska–Lincoln (UNL-CDAS) includes a portable computer, digital timer, tape switches, photocells, and a video camera. A schematic of this system is shown in Figure 1. The portable computer is used to monitor the tape switch and photocell sensors and to record the time of all monitored traffic events. Tape switch sensors are adhered to the road surface at specific locations and used to detect passing vehicles. Photocell sensors are placed over the light-emitting diodes (LEDs) on the load switches (inside the signal controller cabinet) and used to monitor the signal phase change times. The video camera has an in-picture time code capability that makes it an ideal backup data collection system and supplemental visual record of traffic events. In this regard, the videotape record can be used to validate the tape switch data reduction software programs.

Data Gathering and Recording Equipment

The UNL-CDAS uses a 12-MHz 80286 portable computer with a digital-timer board to monitor all system sensors and

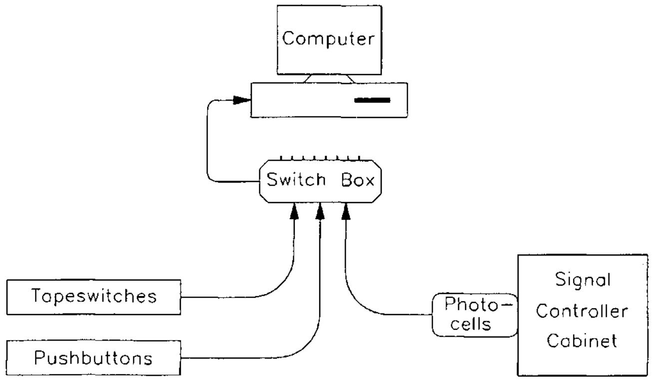

FIGURE 1 Schematic of a computerized data acquisition system.

record traffic event times. The digital-timer board has a 5-MHz clock and 16 digital I/O lines. With this board, the computer can monitor up to 16 sensors. To provide an interface between the computer/board and the sensor lead-in wires, a "switch" box was constructed to act as a terminal block and DC power source for the sensors. The box has a separate switch for each of the 16 sensor inputs to facilitate the mandatory status checks at the start of the data collection process.

Because the timer is a digital device, it produces output in the form of pulses denoting low (closed circuit) and high (open circuit) states. The computer continuously scans the sensors looking for a change in state. During data collection, the computer simultaneously checks the status of all 16 sensors every 0.002 sec. Once this "snapshot" of the switch status is taken, the status signal is decoded, and if a status change detected, the time and line/sensor ID number is recorded to the data file on disk. The computer then makes another status check and the cycle repeats. The scan interval of 0.002 sec results in an event measurement precision of ±0.001 sec, which is adequate for most traffic data needs. Moreover, this scan rate is fast enough to preclude the possibility of missing a switch closure.

Sensing Devices

Two types of sensing devices can be used with a CDAS: tape switch sensors and photocell sensors. Tape switches are used to detect the passage of vehicles at a point on the roadway. Experience with tape switches on urban roadway sections indicates that the life of a typical tape switch is about 70,000 axle hits. However, this lifetime can vary considerably on the basis of weather, pavement, and traffic conditions. Tape switch life appears to be reduced by higher temperatures, coarser pavement surface textures, and severely cracked pavement surfaces. Placement of tape switches on sections of roadway where traffic is accelerating or decelerating (e.g., intersection approaches) can also shorten the life of a tape switch. Tape switches can also experience premature failure because of faulty manufacture or when vehicles with flat tires drive over them. However, such situations are extremely rare.

Photocell sensors can also be monitored by a CDAS. An obvious photocell application would be at a signalized intersection wherein the starting time of the signal phase is needed for purposes of calculating the headway of the first queued vehicle. In this application, photocells can be used to monitor the LEDs on the load switches in the signal controller cabinet. Three LEDs are provided on each load switch, one LED for each signal indication the load switch serves (i.e., green, yellow, and red). In this manner, the status of the signal indications can be continuously monitored and any change (e.g., from red to green) recorded by the computer. This method of monitoring the signal indications prevents the circuitry of the controller from being disturbed in any way, which also has obvious liability benefits.

System Backup

The UNL-CDAS includes a high-resolution video camera to record the same traffic events as the tape switches. This video camera provides 400 lines of horizontal resolution and has an in-picture time coding capability with a measurement precision of ±0.017 sec. The videotape record is used for three reasons:

1. It provides a backup data source in the event of a failure in the tape switch system.
2. It can be used to validate the output of the tape switch data reduction program.
3. In the event of an unusual sequence of time records in the tape switch data, the video can be used to examine visually the sequence of events under question. Occasionally, a multiple-axle vehicle will produce an unexpected sequence of records, or a driver will change lanes without passing over all of the tape switches. In these cases, the data reduction program will flag the unusual sequence and the videotape record will be reviewed for possible explanation.

Types of Data Collected

A CDAS can be used to collect several types of traffic data. The common denominator among these data is that time is a fundamental component (e.g., headway). If two tape switches are located a known distance apart, then characteristics requiring distance and time can be calculated (e.g., speed, acceleration). However, a CDAS is not limited to these characteristics. Vehicle wheelbase and lane position can also be calculated from CDAS data by using the procedures described herein.

Vehicle Headway

Headway data can be collected by placing one tape switch parallel to the direction of travel (i.e., parallel to the vehicle axle) in one traffic lane. Headways are easily determined by computing the difference between the arrival times of the back axles of two successive vehicles. The back axle arrival times can also be used to calculate start-up lost time and saturation flow rate according to the procedures in the 1985 *Highway Capacity Manual* (*1*). If headways are being measured at a signalized intersection, then the load switch LED should be monitored with a photocell to record the start of the signal phase serving the lane under study.

Speed, Acceleration, and Wheelbase

Two parallel tape switches placed relatively close together in the traffic lane can be used to find speed, acceleration, and wheelbase. More specifically, two types of speed can be estimated: average speed and spot speed. Average speed can be found using the following equation:

$$v = \frac{D}{t} \tag{1}$$

where

v = average vehicle speed (m/sec),
D = distance between two parallel tape switches (m), and
t = time it takes one axle to travel across both tape switches (sec)

The calculation of average acceleration and spot speed is based on a procedure described by Evans and Rothery (*2*). This procedure is based on the assumption of a constant acceleration between two parallel tape switches, a distance D apart. The assumption of constant acceleration allows a vehicle's trajectory to be modeled by a second-order equation of distance x as a function of time t. In fact, this assumption allows two equations to be written, one for each axle of the vehicle. This relationship is shown in Figure 2 in terms of the vehicle's trajectory in space and time.

As shown in Figure 2, the trajectories of each axle are identical but separated by a distance equal to the vehicle wheelbase L_w. When a vehicle crosses the trap, it causes four event times to be recorded:

- t_1—the time that the front axle hits the first tape switch,
- t_2—the time that the front axle hits the second tape switch,
- t_3—the time that the rear axle hits the first tape switch, and
- t_4—the time that the rear axle hits the second tape switch.

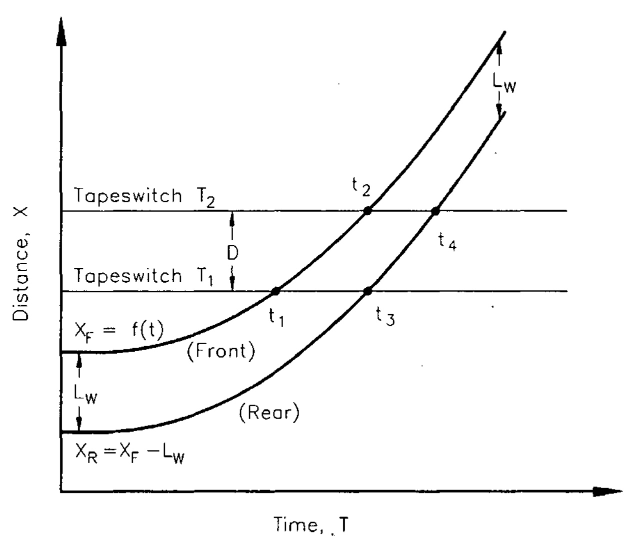

FIGURE 2 Time-space trajectory of vehicle as it traverses two parallel tape switches.

Transforming these event times by subtracting t_1 from each yields the relative travel time t_i' of the front and back axles between the two tape switches. As $x = 0$ when $t_1 = 0$ for the transformed data, the trajectory of the front and back axles can be described by the following second-order equation:

$$x = v_0 t + \frac{1}{2} a t^2 \tag{2}$$

where

x = distance traveled by vehicle in time t (m),
v_0 = velocity of vehicle at instant it hits first tape switch (m/sec),
t = time it takes to travel a distance x (sec), and
a = acceleration of vehicle (m/sec^2).

Examination of the time-space relationships in Figure 2 yields the following relationships for the transformed event times:

$$x = D \quad \text{when } t = t_2'$$
$$x = L_w \quad \text{when } t = t_3'$$
$$x = D + L_w \quad \text{when } t = t_4'$$

where

D = distance between two tape switches (m),
L_w = vehicle wheelbase (m),
T_1 = time when front axle hits first tape switch (sec),
t_2' = time when front axle hits second tape switch less t_1 (sec),
t_3' = time when back axle hits first tape switch less t_1 (sec), and
t_4' = time when back axle hits second tape switch less t_1 (sec).

Substituting these time-space relationships into Equation 2 yields three equations and three unknowns: a, v_0, and L_w. Solving these equations yields the following equations:

$$a = 2\left[\frac{D(t_2' - t_4' + t_3')}{t_2'(t_4' - t_3')(t_4' + t_3' - t_2')}\right] \tag{3}$$

$$v_0 = \frac{D}{t_2'} - \frac{1}{2} a t_2' \tag{4}$$

$$L_w = v_0 t_3' + \frac{1}{2} a (t_3')^2 \tag{5}$$

It should be noted that the equation for v_0, as published in Evans' paper (*2*), is in error; it is shown here in its correct form.

Lane Position

Lane position of vehicles is often a characteristic that is difficult to determine accurately using manual methods. How-

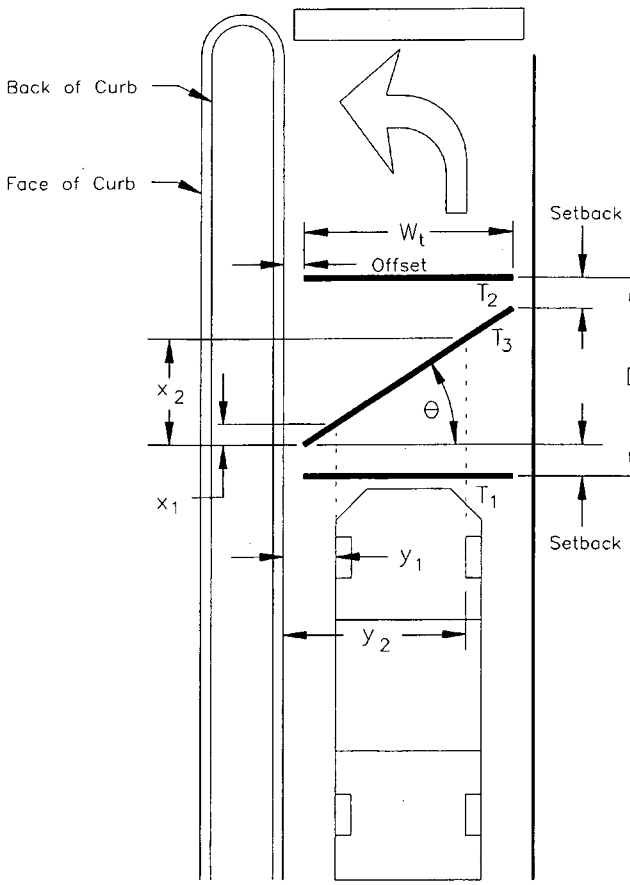

FIGURE 3 Layout of Z-shaped tape switch configuration for estimating lane position.

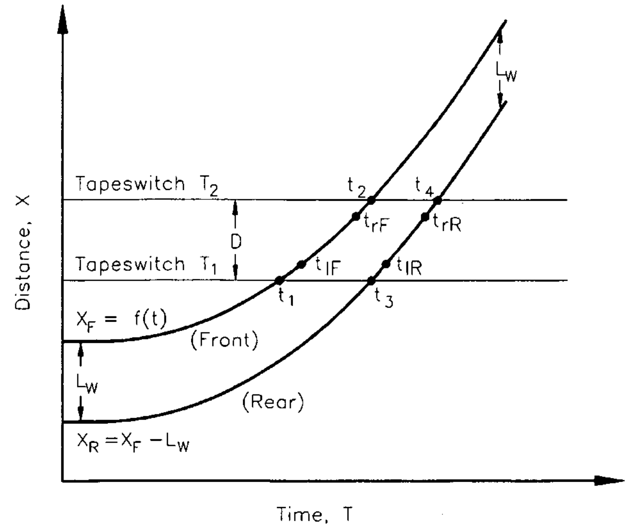

FIGURE 4 Time-space trajectory of vehicle as it traverses Z-configuration.

ever, three tape switches laid out in a Z-configuration can be used to determine lane position more easily and more accurately. Two switches in this configuration are laid parallel in the traffic lane, similarly to the layout for average speed measurement. A third switch is laid diagonally between these two switches. The layout of the Z-configuration is shown in Figure 3. In this particular figure the configuration is shown in a left-turn bay; however, it can also be installed in a through or right-turn lane.

Two assumptions are made in the application of the Z-configuration. The first assumption is that the vehicle has a constant rate of acceleration through the detection zone. This assumption is reasonable, provided that the detection zone is kept relatively short. The second assumption is that the vehicle travels on a relatively straight path through the detection zone. This assumption is reasonable so long as the Z-configuration is not located on a section of roadway where vehicles are executing sharp turning maneuvers.

The method of estimating lane position (i.e., lateral offset from the lane line or curb face) involves the following sequence of calculations. First, the two parallel tape switches are used to estimate the spot speed and average acceleration of the vehicle. Then these characteristics are used to calculate the travel distance, using Equation 2, that coincides with the measured travel time to the diagonal tape switch. Four distances can be estimated in this manner, one for each tire. Finally, using trigonometry, the lateral offset can be calculated from the four travel distances.

The vehicle trajectory through the Z-configuration is shown in Figure 4. This trajectory is identical to that in Figure 2; however, there are four additional event times recorded by the CDAS: t_{lF}, t_{lR}, t_{rF}, t_{rR}. These times coincide with arrival times of the front (F) and rear (R) tires on the left (l) and right (r) sides of the vehicle as it traverses the diagonal tape switch. Transforming these event times by subtracting the event time t_1 from each yields the relative travel time t_i' of the front and back axles through the trap. Having computed the average acceleration and the spot speed from Equations 3 and 4, the travel distance (x_1 + setback) from the first tape switch, T_1, to the point where the left side front tire hits the diagonal tape switch, T_3, can be calculated as

$$x_1 + \text{setback} = v_0 t_{lF}' + \tfrac{1}{2} a (t_{lF}')^2 \tag{6}$$

where t_{lF}' is the left front tire travel time between T_1 and T_3 in seconds and setback is the distance between T_1 and the nearest end of T_3. The value of x_1 is calculated from the known value of setback (as determined from the Z-configuration's layout in the traffic lane).

The lateral offset distance from the face of curb to the point at which the driver's side front tire hits T_3 can be calculated as

$$y_1 = \frac{x_1}{\tan\theta} + \text{offset} \tag{7}$$

with

$$\theta = \tan^{-1}\left(\frac{D - 2\ \text{setback}}{W_t}\right) \tag{8}$$

where

y_1 = lateral offset distance to left front tire (m),
θ = angle between T_3 and a line parallel to T_1,
W_t = width of tape switch sensor (m), and
offset = lateral distance between face of curb and edge of tape switch trap (m).

In a similar manner, Equations 6, 7, and 8 can be used to calculate the lateral offset to each of the remaining three vehicle tires.

Once the lateral offsets to the tires have been calculated, the lateral offset to the vehicle centerline of each axle can be obtained by taking an average of these tire offsets. One problem with this approach is that the tire offsets must be corrected for the width of the tire contact patch. This can be estimated at about 15 cm (6 in) for most front axle tires. But the large number of trucks with dual tires on the rear axle makes estimation of the tire width (and rear axle lane position) more difficult. As a result, it is recommended that only the front tires be used in estimating the vehicle centerline offset. This offset distance (i.e., lane position) can be calculated as

$$L_p = \frac{y_1 + y_2 + P_w}{2} \tag{9}$$

where L_p is the distance in meters from the face of the curb to the centerline of the vehicle and P_w is the average width of a tire patch (0.15 m, 0.5 ft).

Phase Duration and Cycle Length

Information about the duration of various intervals in the signal cycle are sometimes a necessary part of the data collection effort. With a photocell sensor, the duration of the green, yellow, and red phase intervals as well as the cycle length can be measured accurately by the CDAS. As mentioned previously, the photocells would be placed on the LEDs on the load switches in the controller cabinet. The event times corresponding to a change in phase interval can then be used to calculate the duration of the desired interval by simple subtraction.

FIELD STUDY CONSIDERATIONS

Advance Preparation

Most of the preparation for the study is performed well before the study date. Plan sheets of the study sites are obtained and used to help in estimating the amount of wire and adhesive material needed for sensor installation. The plan sheet is also used to determine tape switch locations and the position of the signal controller cabinet. The CDAS data recording equipment are typically set up on the same corner as the signal controller cabinet to minimize the total length of lead-in wire. An example of a tape switch layout for determining the lane position of left-turn vehicles and headways of through vehicles is shown in Figure 5.

Special materials are needed to protect the tape switches from the harsh environment of the roadway. One of these

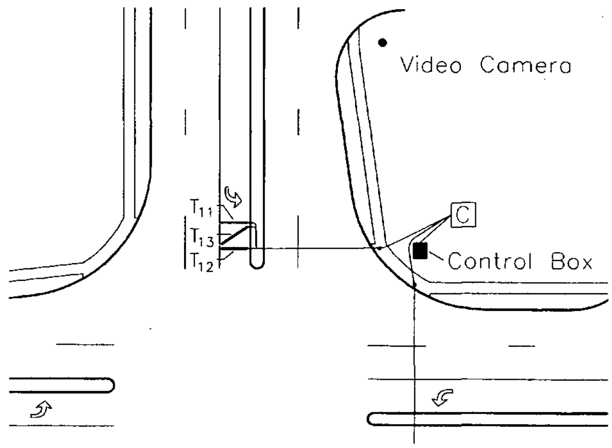

FIGURE 5 Intersection tape switch layout for lane position and headway estimation.

materials is a 5.1-cm (2-in.) felt pad strip laid beneath the 1.6-cm (⅝-in.) tape switch. This pad protects the tape switch from irregularities in the pavement surface. Experience has also shown that a tape switch installed without a pad will slide back and forth with each passing vehicle. Under higher traffic demands, this will produce a sanding action that will eventually wear the bottom off of the switch and cause it to fail prematurely.

The tape switch and pad are adhered to the road surface with a 10-cm (4-in.) bituthene mat material. This material has been found to have excellent adhesive properties under a wide range of temperature conditions. Moreover, the fiberglass in the mat is able to withstand abrasion from frequent tire impacts and thereby protects the tape switch for long periods (generally much longer than the life of the tape switch).

Tape switch installation proceeds one lane at a time. This procedure has been found to yield the safest conditions for the installation crew. It also minimizes the impact on traffic by requiring only one lane to be closed at a time. To facilitate this approach, both the mat and pad are prepared in 3.7-m (12-ft) lengths.

The lead-ins used with the UNL-CDAS are four-conductor 20 AWG wire. Use of larger wire gauge is not recommended as it tends to incur a larger impact from the crossing vehicle tire. Moreover, in some instances, a tape switch and lead-in wire may share the same pad. In this situation, a larger wire may deflect the tire over the tape switch and result in an intermittent loss in data.

The width of the tape switch trap must be determined in advance and is dependent on the type of data to be collected. For a headway study, the tape switch should be centered in the traffic lane with about 0.30 m (1 ft) of clearance to the adjacent lanes. This approach will eliminate errant hits on the switch by vehicles that are effectively in the adjacent lanes (but that may wander momentarily out of those lanes). On the other hand, tape switches for a lane position study should be long enough to span the full width of the lane. As noted in the previous section, both the left and right side tires must be detected by the diagonal tape switch. Thus, without full lane coverage, it would be possible for vehicles traveling close to the edge of the lane to traverse the Z-configuration without the necessary four tires being detected. These extreme vehicles may be of particular interest to the study and cannot be missed.

The location of the tape switch on an intersection approach lane is dependent on the type of data to be collected. Since the equations for predicting lane position assume a straight-line travel path, it is important that the Z-configuration be placed in advance of the stop line. For speed or acceleration measurements, the trap can be placed in advance of or beyond the stop line as needed. For headway measurement, the tape switch should be placed just beyond the line behind which vehicles consistently stop (e.g., the crosswalk or stop line). In all cases, the tape switches must be perpendicular to the direction of traffic flow.

Z-Configuration Layout

The layout of the Z-configuration is dictated by several physical constraints, which include width of lane, proximity of Z-configuration to the stop line, and length of detection zone. An example Z-configuration layout is shown in Figure 5.

In general, the derivation of the speed, acceleration, and lane position equations is based on the assumption of constant acceleration between the two parallel tape switches. This assumption becomes less valid as the distance between the two parallel tape switches increases. Therefore, this distance should not be much larger than the wheelbase of a passenger car.

The effect of this distance on the angle between the diagonal tape switch and the first parallel tape switch should also be considered. For extremely short distances, the angle will be quite small and tend to magnify (via Equation 7) any error in the estimate of x_1. This error would then be reflected in the estimate of lane position. Therefore, the distance between the two parallel tape switches should not be much shorter than the wheelbase of a passenger car.

On the basis of these considerations, it was determined that the distance between the two parallel tape switches should be 3.0 m (10 ft). This distance represents a good compromise solution and does not appear to have any inherent problems. In fact, this distance is consistent with that used by Evans and Rothery (*2*).

The setback distance between the parallel tape switches and the near edge of the diagonal tape switch (Figure 3) was established as 0.46 m (1.5 ft) for somewhat more subjective reasons. The main goal was to avoid having the ends of two tape switches under a tire at the same time because it could cause the tire to vault over the second switch. Thus, it was determined that the setback distance should exceed the length of a tire contact patch.

Tape Switch Installation

The actual installation should begin at the point farthest from the curb (i.e., where tape switch is located) and proceed toward the curb. The tape switch or lead-in wire should be installed in one lane at a time to minimize the delay to motorists. In general, an experienced installation team will need about 1 min to install a tape switch or lead wire in a traffic lane. This time will increase in situations where multiple lead-ins are being "carried" to the curb. In this regard, each additional lead wire will add about 30 sec to the installation time in a traffic lane.

A tape switch installation crew usually has four to six members. This crew is composed of three or four installers and one or two flaggers. Three installers and a flagger are adequate for smaller studies; however, four installers and two flaggers are desirable for larger studies requiring multiple lead-in wires in a traffic lane. Two flaggers will maximize the conspicuity of the work area and provide positive driver guidance around the installation crew. These flaggers should be supplemented with one or two arrow-board trucks operated by local agency staff. These trucks can be brought into position sequentially in the appropriate lane (near or far side of the intersection) and used to provide advance warning of the closure of that particular lane.

If both a tape switch and wire are being installed under the same mat adhesive, the tape switch should be positioned nearest to the oncoming traffic. In other words, the wire lead-ins from other tape switches should be placed behind the tape switch on the pad. The reason for this arrangement is to ensure that the tape switch is the first point of vehicle-tire contact. If the wire leads are placed before the tape switch on the pad, the wire may deflect the tire over the tape switch and result in lost data. Precautions should be taken to make sure that multiple wire lead-ins are placed parallel to each other on the pad. If these wires are allowed to cross (or overlap) one another, there is a possibility that one wire will cut into the other wire under the weight of a crossing vehicle.

QUALITY CONTROL

Several measures are taken to ensure the quality of the collected data with the UNL-CDAS, including backup systems and formalized procedures for monitoring sensor performance. As mentioned previously, a video camera with an in-picture time-code capability is used as a backup data collection system. This camera is located such that all points of interest can be captured on the videotape image. When one video camera is not able to capture all desired events, a second video camera is used.

During data collection, the performance of the CDAS is monitored continuously via the computer display. This monitoring is essential to avoid losing large amounts of data due to equipment failure. The computer display provides a status report of each sensor and lead-in. Should a failure be detected, its time of occurrence is noted in a log book and steps are taken to correct it.

In addition to the continuous monitoring of the system status, a visual check of the tape switches and mat adhesive conditions is performed periodically. Occasionally, a piece of mat material will be damaged by a spinning or sliding tire and will require a mat overlay. In general, this check is performed at least once an hour during the study, although the frequency may be increased to every ½ hr if the temperature is warm or traffic demands are high.

COST

Although a CDAS requires the use of relatively expensive technology, it remains a practical and cost-effective method

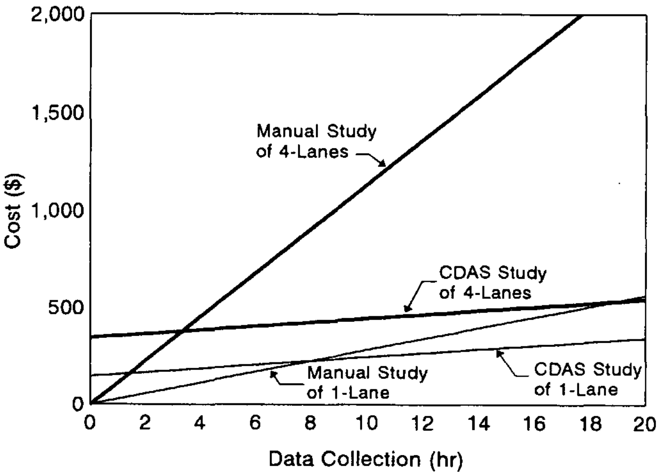

FIGURE 6 Cost comparison for CDAS and manual methods for headway study.

of collecting data. In fact, when compared with manual methods, a CDAS can save time and money. To illustrate this point, the costs of conducting one- and four-lane headway studies using both the CDAS and manual methods were done. The results of this comparison are shown in Figure 6.

The costs considered for this comparison included the startup costs (i.e., materials plus labor for installation) and ongoing costs (i.e., replacement materials and labor during study). The CDAS startup costs include the tape switches, mat, and pad as well as the setup time for a crew of five. The CDAS ongoing costs reflect the labor cost for one person to monitor the system. The manual method does not have any significant startup costs but does incur considerable ongoing labor costs (which vary with the size and duration of the study).

The trend shown in Figure 6 suggests that the manual method costs more than the CDAS method for studies of even nominal duration. The cost of the manual method is lower than CDAS only for studies of short duration. This trend stems from the higher initial costs of the CDAS equipment. As the duration of the study increases, the higher labor costs ($10/hr) of the manual method offset the higher initial costs of the CDAS method. For a one-lane headway study, the manual method has a lower cost for a study of fewer than 8 hr, but for a study of more than 8 hr, the CDAS method has the lower cost. For a four-lane headway study, the break-even point is reached at about 4 hr because of the higher on-going labor costs of the manual method.

VALIDATION

The UNL-CDAS was validated by comparing it with two other data collection techniques. The first technique used a verified videotape system with a built-in time coding capability. The second technique used a stopwatch to measure event times in a more traditional manner.

The videotape system represented the more accurate validation technique because it records to a precision of ± 0.017 sec. Fifty headways were measured and compared using both the videotape system and UNL-CDAS. The average difference between the headways measured with the videotape system and those measured with the CDAS was 0.005 sec. This difference was not significantly different than 0 (at a 95 percent level of confidence).

The UNL-CDAS was also validated using a manual technique to measure headways. For this test, two observers collected headway data for a continuous period of 2 hr. Again, a comparison of the observed and CDAS headways indicated no significant differences.

During the second test, several interesting observations were made about the ability of the observers to collect data. As time progressed and the observers tired, they tended to underestimate vehicle headways. Moreover, the magnitude of this underestimate increased with time, as did the standard deviation of the error (i.e., the difference between the observed and CDAS headways). Given that a study of vehicle headways typically requires hundreds and perhaps thousands of observations for statistical significance, it appears that automated data collection systems offer the only realistic means of measuring headways (and other traffic characteristics) with any reasonable degree of accuracy.

CONCLUSION

A computerized data acquisition system using tape switch and photocell sensors can be an effective way of accurately collecting a variety of traffic characteristics. To conduct a successful computerized tape switch study requires a great deal of advance planning; however, the computerized study also offers a great deal of flexibility and can be an efficient means of obtaining a large amount of traffic data. Many types of driver characteristics such as startup lost time, clearance lost time, and lane position can be measured.

The computerized data acquisition system was validated and found to be accurate and precise in its measurement of timed events. The accuracy and precision of the computerized data acquisition system is its most important advantage over manual studies. As a result, small differences in driver behavior can be more easily detected and the statistical significance of these differences more easily established. Another advantage of a computerized data acquisition system is that it typically takes less time to collect and enter the data than manual methods.

REFERENCES

1. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985
2. Evans, L., and R. W. Rothery. Influence of Vehicle Size and Performance on Intersection Saturation Flow. *Proc., 8th International Symposium on Transportation and Traffic Theory*, University of Toronto Press, Canada, 1981, pp. 193–222.

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

Using Video Technology To Conduct 1991 Boston Region External Cordon Survey

KENNETH S. MILLER, THOMAS N. HARVEY, PAUL W. SHULDINER, AND CECILIA W. HO

The 1991 Boston Region External Cordon Survey is described. External cordon surveys are designed to obtain information on trips that cross the external boundary of a study area. The primary use of external survey data is the development and calibration of external trip distribution models. Particular emphasis is given to two innovative aspects of the 1991 Boston survey: (a) the application of small sample design techniques for the efficient design of an appropriate sample in order to gather significant data on variables of interest, and (b) the extensive use of video technology. When conducting roadside interviews or handing out postcards is not possible because of restrictions regarding traffic delays, the study has shown that video can be used successfully to obtain the needed information, using off-the-shelf equipment and specific techniques to minimize costs.

In 1990 the Central Transportation Planning Staff (CTPS), the staff to the metropolitan planning organization (MPO) in the Boston region, initiated the Regional Planning Study (RPS), a multiyear project to update the model set for estimating travel demand in the Boston region. The project included an extensive data collection effort designed to provide the requisite information for model development.

This paper describes the 1991 Boston Region External Cordon Survey, one component of the RPS data collection effort. External cordon surveys are designed to gather information on trips that cross the external boundary of a study area. These include external-external (E/E) trips, which begin and end outside the study area but travel through it; internal-external (I/E) trips, which begin inside the study area but end outside; and external-internal (E/I) trips, which originate outside the study area but end within it. Vehicles crossing the cordon line are the basic sampling unit, for which the following basic travel data are usually collected: trip origin and destination by address or nearest intersection, travel time, trip purpose, vehicle occupancy, and vehicle class.

The primary use of external survey data is the development and calibration of external trip distribution models. Origin-destination (O-D) survey data are used to develop estimates of average trip lengths so that trip-length frequency distributions (TLFDs) can be used for model calibration. This is done because direct estimation of external vehicle trip tables from survey O-D data is problematical: survey data at the required zonal level are statistically insignificant at useful confidence and precision levels. In addition, data from the external survey can also be compared with information from the 1990 census and the 1991 Boston Region Household-Based Survey to analyze the relationship among demographics, economic activities, and travel behavior.

This paper summarizes the available techniques for conducting an external cordon survey and describes the methods used for the study. Particular emphasis is given to two innovative aspects of this survey: the application of small-scale sample design techniques and the extensive use of video technology.

AVAILABLE SURVEY METHODS

External cordon surveys are usually conducted using one of three methods: roadside interview, roadside postcard survey distribution (postcards are to be mailed back), or license plate recording/matching with a survey mailed out to be returned (*1*).

Roadside Interview

Perhaps the most common method for gathering the requisite data is to administer a roadside oral interview at the selected survey stations. This method provides the needed data quickly and allows the interviewer to prompt and probe for complete information. In addition, the response rate is usually higher than for the other methods, reducing nonresponse biases. Because there is no delay in estimating the response rate, an adequate sampling rate can be used to achieve the goals of the sample design. No mailing costs are incurred, so costs tend to be lower than for other methods, although this must be balanced against the probable higher personnel costs.

Several possible disadvantages should be noted. Conduct of the survey may cause traffic delays, particularly during the peak period and at high-volume locations. Management of the survey may be complicated because of the need to involve several organizations, including the police and the highway department. Safety of the surveyors may also be an issue.

Postcard Surveys

Another method is to stop vehicles at the survey station and hand out a postcard to be mailed back after completion by

K. S. Miller, Central Transportation Staff, 10 Park Plaza, Boston, Mass., 01226; current affiliation: Bureau of Transportation Planning and Development, Executive Office of Transportation and Construction, Commonwealth of Massachusetts, 10 Park Plaza, Boston, Mass. 02116. T. N. Harvey, HMM Associates, Inc., 396 Baker Avenue, Concord, Mass. 01742; current affiliation Cambridge Systematics, Inc., 222 Third Street, Cambridge, Mass. 02142. P. W. Shuldiner, Department of Civil Engineering, University of Massachusetts, Amherst, Mass. 01003. C. W. Ho, Central Transportation Planning Staff, 10 Park Plaza, Boston, Mass., 02116; current affiliation: JHK and Associates, 4660 Kenmore Avenue, Alexandria, Va. 22304.

the respondent. This procedure is somewhat less likely to disrupt traffic than conducting interviews is, and it requires fewer field personnel. However, the response rate is usually lower, in the range of 20 to 30 percent. The low response rate may result in significant nonresponse bias. Costs may be significantly higher than for a roadside interview, also: a recent study in San Antonio that compared roadside interviews with postcard distribution concluded that the unit cost of valid responses was several times higher for the postcard method than for the roadside interview method (*2*).

License Plate Recording and Survey Mailout

A third method involves recording license plate numbers as vehicles pass, matching plate numbers against registry files to determine home addresses, and mailing a survey form to be returned by mail. The license plates can be recorded by using pencil and notepad, audiotape recorder, laptop or portable computer, or film or video camera. License plate matching is usually done through the use of a computerized matching procedure. The major advantage of the license plate method is that it does not disrupt traffic.

There are several disadvantages to the license plate approach, however. The lag between time of observation and the receipt of the survey by the respondent may lead to low response rates (and significant nonresponse bias) and high recollection errors. The method can be fairly complicated to manage, and costs may be higher than for the other methods. Adverse publicity may be generated because of errors in license plate recording or address matching and the subsequent mailing of surveys to the wrong household. Even if errors are kept to a minimum, the public may perceive it as an invasion of privacy.

The license plate method was selected for the Boston survey for one overriding reason: the authors were not allowed to disrupt or delay traffic flow in any manner, at any location.

Supplemental Information

In addition to the actual survey data, two types of vehicle count information are needed to weight and expand the sample data to represent the entire vehicle population for the study area:

1. Manual vehicle counts at survey stations to obtain information on vehicle classification and vehicle occupancy during the survey.
2. Automatic traffic recorder (ATR) counts to collect average weekday hourly traffic information for both travel directions at each survey station.

SAMPLE DESIGN

The external cordon line of the Boston RPS area is shown in Figure 1. At the cordon line, 102 roadways were identified as possible survey sites. It was estimated that 80 percent of the daily traffic crosses the cordon through the 33 highest-volume stations, 85 percent through 40 stations, and 90 percent through 47 stations. Because of budgetary limitations, and for purposes of efficiency, it was decided that traffic at the 47 stations with an average daily traffic (ADT) greater than about 5,000 should be surveyed, capturing 90 percent of the traffic crossing the cordon. To examine the TLFDs of facilities with low traffic volumes, three stations with ADTs of less than 5,000 were also included. This selection resulted in 50 external stations at which surveys were performed.

Previous surveys of this type have relied on one of two approaches to sampling. The first approach focuses on providing a sufficient sample size for direct estimation of traffic flows 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 trip purpose or vehicle classification (*3*) or as it relates to a particular destination (*4*). Neither approach correlates explicitly with the primary purpose of these kinds of surveys: the estimation of zonal-interchange-level external trip tables.

Unfortunately, direct estimation of trip tables from survey data at a useful statistical level of confidence and precision is problematical: for a large region, with perhaps a thousand traffic zones (and consequently a million possible trip interchanges), it is almost impossible to collect enough data for direct estimation of trip tables (even recognizing that a proportion of trip interchanges will have zero trips). An alternative to the use of survey data for direct trip table estimation is the use of survey data to develop estimates of parameters (particularly TLFD) for calibration of external trip distribution models. The distribution models are then used to estimate zonal-interchange external trip tables. (Of course, statistically significant aggregate interchange trip data are useful for validating the distribution models.)

Therefore, for this survey, trip length was chosen as the variable of interest. Both Smith's (*5*) and Stopher's (*6*) small-sample survey design methods were applied to determine the required sample size for each external station in order to ensure a statistically significant estimate of average trip length, the variable of interest, at a 90 percent confidence level and a ± 5 percent precision level. In essence, small-sample design procedures allow the use of previously collected information about the variance of the variable of interest to efficiently design a sampling procedure to collect updated information about those variables.

The following procedures were used to apply the small-sample design techniques:

1. The survey stations were grouped into different geographic sectors corresponding to the CTPS sectors. Since the study area is in the shape of a semicircle and the major roadways are in a radial pattern, differences in the TLFD among major roadways can be identified. All the external survey stations are included in one of the six sectors: North Shore, North, Northwest, West, Southwest, and South.
2. The proportions of I/E, E/I, and E/E trips to total external trips were extracted from the existing CTPS regional highway trip table, which is based on previous survey data. Estimates of mean trip length and trip length coefficient of variation for each survey station were also obtained from the regional trip table. An estimate of the number of completed and valid surveys was then calculated for each survey station by applying Smith's and Stopher's sample design procedures.

It should be noted that the estimates of trip length and trip length coefficient of variation developed from existing infor-

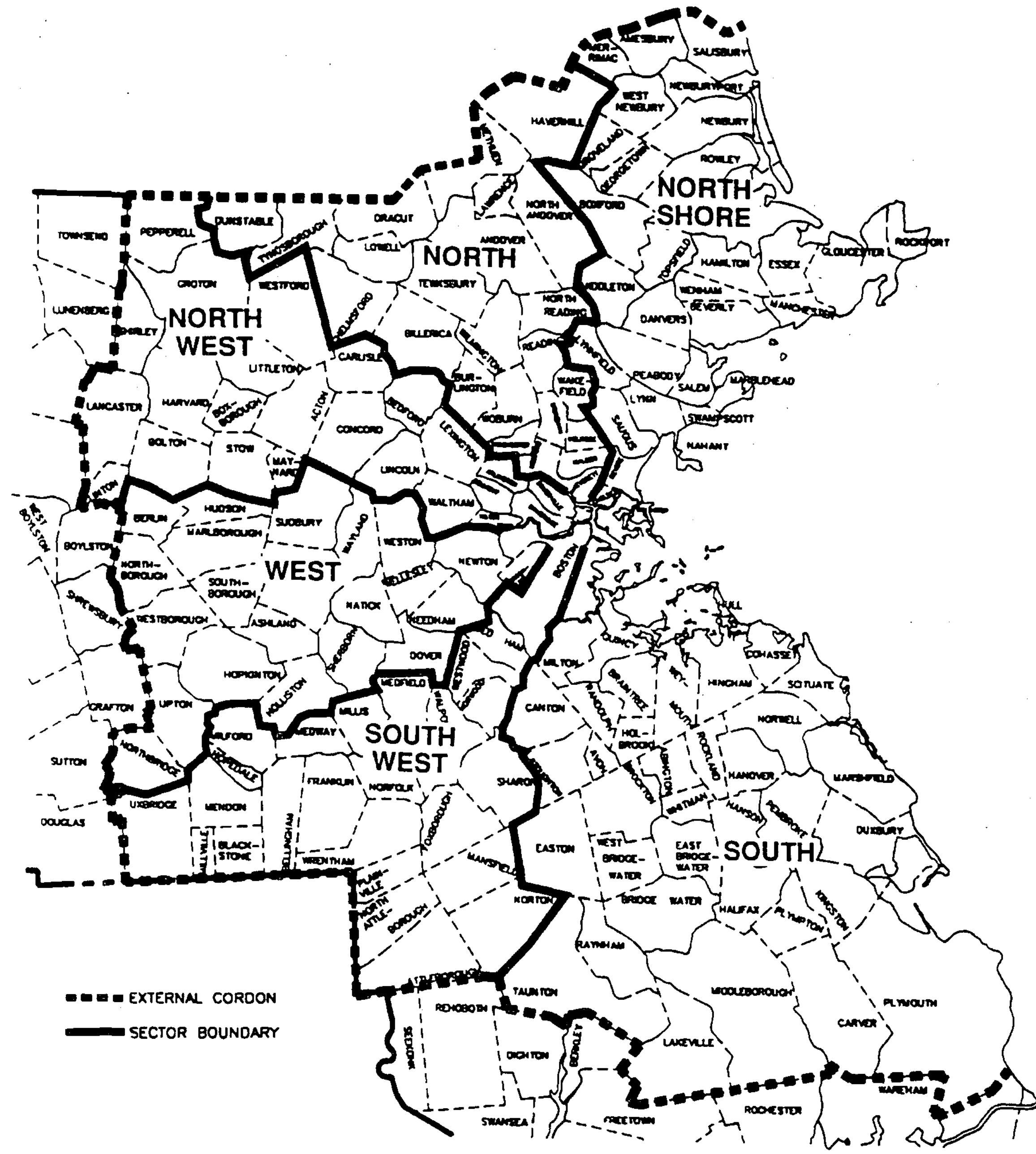

FIGURE 1 Boston region external cordon line and sector boundaries.

mation may not be accurate or up to date, which in turn may result in inaccuracies and inefficiencies in the sample design. This is inherent in the application of small sample design techniques; after all, replacing outdated data is the reason for conducting surveys. An examination of the errors introduced by using the existing data for sample design can be conducted after the new survey is analyzed and compared with sample design input assumptions.

3. On the basis of the sample size estimates from the previous step, the survey stations were then classified into three different groups according to their 1987 ADT:

–Low-volume stations (ADT less than 10,000),

–Medium-volume stations (ADT between 10,000 and 35,000), and

–High-volume stations (ADT higher than 35,000).

The station at which the highest proportion of traffic needed to be surveyed was chosen as the critical station for each ADT group. The number of completed and valid surveys required at each station was generated by applying the critical cell proportion to each station in the ADT group. This was done because it is difficult to design a survey method to survey the exact number of samples needed at each station.

The sample design procedure resulted in the following required sampling proportions for the three ADT groups:

- Low-volume stations: 10 percent of the inbound traffic crossing the cordon during the survey period.
- Medium-volume stations: 6 percent of the inbound traffic crossing the cordon during the survey period.

• High-volume stations: 5 percent of the inbound traffic crossing the cordon during the survey period.

These percentages represent the final proportion of traffic for which completed and valid surveys are needed. To generate an estimate of the required number of complete and valid surveys, the number of license plates that needed to be recorded was estimated by adjusting for the following factors:

• Plate recording and matching error: it was assumed that 85 percent of recorded plates could be matched with the registry files to obtain the home addresses of the respondents.
• Nonresponse: it was assumed that 25 percent of the surveys mailed would be returned. This takes into account surveys not delivered because of inaccurate address matching, errors in matching, and nonresponse.
• Incomplete surveys: it was assumed that 95 percent of the returned surveys would be complete.

Because of the combined effect of these assumed rates, it was estimated that surveys would be returned for only 20 percent of the recorded plates (85 percent × 25 percent × 95 percent × 20 percent). Table 1 presents an estimate of the number of complete and valid surveys needed at each station, on the basis of estimated 1987 ADT and appropriate non-passenger-vehicle factors; the estimated total regional sample size was 24,900 complete and valid surveys. It should be noted that the actual required number of completed surveys was related to actual traffic volumes recorded during the survey process as well as adjustments for the actual rates for the preceding factors.

SURVEY PROCEDURE

Survey Schedule

The survey was performed in two stages: the first extended from May 7 to June 13, 1991; the second, from September 10 to October 22, 1991. Each station was surveyed for 1 day from 6:00 a.m. to 7:00 p.m. on Tuesdays, Wednesdays, or Thursdays during the survey period. Generally two sites were surveyed per day.

License plate recording and vehicle occupancy and classification counting were conducted for inbound traffic only, and 24-hr ATR counts were conducted in both directions for the full midweek period for each station during its survey week. An observer alternated counting classification and occupancy at 15-min intervals during three 2-hr survey periods: a.m., p.m., and midday. Two types of recording equipment were used for the occupancy and classification counts: denominator boards were used at some locations, and intersection boards programmed for occupancy and classification were used at other locations.

The surveys were mailed within 5 days of plate recording so that the prospective respondent would receive the form no later than 1 week after being observed. Several tasks had to be performed within that 5-day period: transcription and key entry of the plate data, matching with vehicle registry files to obtain home addresses, printing of vehicle-specific data on the forms, and form folding and sealing. Up to 10,000 surveys were mailed each week during the survey period.

Survey Instruments

Three survey forms were required: a mailed survey questionnaire, a vehicle occupancy count form, and a vehicle classification count form. The mailed survey form had to meet several criteria:

• Each form had to contain observation-specific information: observed license plate, as well as the location, direction, date, and time of observation. This was accomplished by preprinting forms that could be form-fed to printers to receive the observation-specific data during the mail-merge process.
• To keep postage costs at a minimum, it was decided to conform to U.S. Postal Service postcard maximum size requirements while ensuring room for necessary survey questions, respondent and return addresses, and appropriate cover information and instructions.
• The paper stock had to be heavy enough to withstand the printing, mail-merge, and mailback processes, but not so heavy that it could not be form-fed.

The survey form, shown in Figure 2, was folded in half before mailing; the respondent then tore off and returned the bottom half of the form. A trifold form was considered and rejected because of the additional folding costs.

License Plate Recording Methods

Several methods are available to record license plate data for surveys of this kind. The most common methods, used in previous studies, include the use of notepad and pencil, audiocassette recorders, laptop computers, and video. Certain aspects of this survey were important to consider in the selection of appropriate plate recording methods:

• It was expected that a significant number of plates would be recorded from Massachusetts, New Hampshire, Rhode Island, and Maine.
• For each state, there are at least 40 plate formats in use for passenger vehicles.
• Massachusetts requires only rear plates on passenger vehicles; therefore, the methods selected must work for reading rear plates.
• In Massachusetts all pickup trucks and vans are required to have commercial plates even if they are used as private passenger vehicles.
• Many of the stations do not have available overhead locations, so plate recording must be conducted from the side of the road.
• The daily traffic volumes at the 50 survey stations vary from about 5,000 to more than 80,000 ADT.

Field-crew training provided the initial means for assessing each method; a pilot survey was then used to further evaluate the methods. Three stations were used in the pilot survey to test the plate recording methods: (a) notepad, audiocassette recorder, and laptop computer were tested at a low-volume station; (b) audio and video were tested at a medium-volume station; and (c) video was tested at a high-volume station. Particular attention was paid to determining the capabilities of crew members with respect to each of the methods. The

TABLE 1 Estimated Sample Size for Each Survey Station

Station	Town	Location	Estimated 1987 ADT	Station Volume Class	Estimated Inbound Traffic 6AM-7PM	Estimated Required Completed Surveys	Matched Plates Required at 22.5% Response	Estimated Plates Recorded at 85% Match Rate
1	Salisbury	I-95 @ N.H. SL	62,000	H	25,110	1,252	5,567	6,549
2	Salisbury	Rt 1 @ N.H. SL	16,800	M	6,804	408	1,814	2,135
3	Salisbury	Rt 1A @ N.H. SL	10,500	M	4,253	255	1,134	1,334
4	Salisbury	Rt 286 E of Rt 1	15,200	M	6,156	368	1,636	1,925
5	Salisbury	Main St @ N.H. SL	6,300	L	2,552	255	1,134	1,334
6	Amesbury	Rt 150 @ N.H. SL	9,700	L	3,929	393	1,746	2,054
7	Haverhill	Rt 108 @ N.H. SL	6,400	L	2,592	259	1,152	1,355
8	Haverhill	Rt 125 @ N.H. SL	32,800	M	13,284	797	3,542	4,168
9	Haverhill	North Ave @ N.H. SL	12,700	M	5,144	309	1,372	1,614
10	Haverhill	Rt 121 @ N.H. SL	7,400	L	2,997	300	1,332	1,567
11	Methuen	I-93 @ N.H. SL	80,000	H	32,400	1,620	7,200	8,471
12	Methuen	Rt 97 @ N.H. SL	8,200	L	3,321	332	1,476	1,736
13	Methuen	Rt 28 @ N.H. SL	25,300	M	10,247	615	2,732	3,215
14	Dracut	Rt 38 @ N.H. SL	12,300	M	4,982	299	1,328	1,563
15	Dracut	Mammoth Rd @ N.H. SL	11,500	M	4,658	279	1,242	1,461
16	Tyngsborough	Middlesex Rd S of Rt 3	16,000	M	6,480	389	1,728	2,033
17	Tyngsborough	Rt 3 @ N.H. SL (S of x36)	62,000	H	25,110	1,256	5,580	6,565
18	Pepperell	Rt 119 @ Townsend TL	12,400	M	5,022	301	1,339	1,576
19	Lunenburg	Rt 2A W of 2A & 225	9,000	L	3,645	365	1,620	1,906
20	Lancaster	Rt 2 @ Leominster TL	43,000	H	17,415	871	3,870	4,553
21	Leominster	Rt 117 W of Lancaster TL	5,300	L	2,147	215	954	1,122
22	Sterling	Rt 62 W of Lancaster TL	5,000	L	2,025	203	900	1,059
23	Clinton	Rt 70 @ Boylston TL	6,300	L	2,552	255	1,134	1,334
24	Northborough	I-290 @ Shrewsbury TL	55,200	H	22,356	1,118	4,968	5,845
25	Shrewsbury	Main St @ Northboro TL	7,500	L	3,038	305	1,354	1,592
26	Shrewsbury	Rt 9 @ Northboro TL	36,000	H	14,580	729	3,240	3,812
27	Shrewsbury	Rt 20 S of N TL & Rt9	20,000	M	8,100	486	2,160	2,541
28	Grafton	I-90 btwn Exit 11 & 11A	57,600	H	23,328	1,166	5,184	6,099
29	Upton	Rt 140 @ Grafton TL	8,200	L	3,321	332	1,476	1,736
30	Northbridge	Rt 122 @ Grafton TL	8,300	L	3,362	334	1,485	1,747
31	Bellingham	Rt 126 S of S Main St	14,700	M	5,954	357	1,588	1,868
32	Uxbridge	Rt 16 E of Rt 146	5,200	L	2,106	211	936	1,101
33	Uxbridge	Rt 146A E of Rt 146	10,000	M	4,050	243	1,080	1,271
34	Wrentham	Rt 121 @ R.I. SL	5,600	L	2,268	227	1,008	1,186
35	N Attleboro	I-295 @ R.I. SL	25,000	H	10,125	505	2,246	2,642
36	Attleboro	County St @ R.I. SL	7,800	L	3,159	316	1,404	1,652
37	Attleboro	Rt 123 W of Rt 1	11,200	M	4,536	271	1,204	1,417
38	Attleboro	I-95 @ R.I. SL	72,000	H	29,160	1,458	6,480	7,624
39	Attleboro	Rt 1 @ R.I. SL	20,000	M	8,100	486	2,160	2,541
40	Attleboro	Rt 1A @ R.I. SL	24,000	M	9,720	583	2,592	3,049
41	Attleboro	Rt 152 W of Thurber Ave	5,600	L	2,268	225	999	1,175
42	Rehoboth	Rt 118 @ Attleboro CL	6,900	L	2,795	279	1,242	1,461
43	Taunton	Rt 44 btwn N & S Walker	15,000	M	6,075	365	1,620	1,906
44	Dighton	Rt 138 N of Center St	8,300	L	3,362	336	1,494	1,758
45	Taunton	Rt 24 S of Rt 140	43,000	H	17,415	871	3,870	4,553
46	Taunton	Rt 79 @Taunton/Berk CL	5,000	L	2,025	203	900	1,059
47	Lakeville	Rt 140 @ Freetown TL	20,000	M	8,100	486	2,160	2,541
48	Lakeville	Rt 105 @ Rochester TL	2,500	L	1,013	101	450	529
49	Middleboro	I-495 @ Rochester TL	32,000	H	12,960	648	2,880	3,388
50	Plymouth	Rt 3 1.5 mls N of Bourne	27,600	M	11,178	671	2,981	3,507
		TOTAL:	1,030,300		417,300	24,900	110,700	130,200

traffic volume and speed ranges over which each method might be used by qualified crew members were estimated. The following summarizes the evaluation of each method.

Notepad

The notepad method demonstrated that one person could record as many as 600 license plates per hour under ideal conditions; achieving the ideal required a steady stream of vehicles so that they could be recorded at 5-sec intervals. A well-disciplined recorder using a notepad could do this by recording every other vehicle where the headways approximated 2 to 3 sec on a steady basis or even in platoons. Recording accuracy varied greatly among the field personnel, however. Some difficulties occurred due to violation of the minimum headway limit and efforts of the recorder to record too many numbers in a given period.

It was recognized that ideal conditions probably would not occur during the main survey. Therefore, it was decided that the notepad method would be appropriate only at low-volume stations in the main survey.

COMMONWEALTH OF MASSACHUSETTS
DEPARTMENT OF PUBLIC WORKS

A Message from the Secretary of Transportation

The Massachusetts Department of Public Works, in cooperation with the Federal Highway Administration, is conducting a study to improve the traffic conditions in the Eastern Massachusetts area. We need your help and opinions in order to plan improvements effectively.

Your answers to this survey are completely confidential. The success of this timely and important survey depends heavily on your participation. Please take a few minutes to answer the questions on the attached survey form. Your prompt completion and return of this questionaire will be greatly appreciated.

Sincerely,

Richard L. Taylor
Secretary of Transportation

YOUR INFORMATION COUNTS

No matter how much or how little you travel, **YOU ARE IMPORTANT.** You are one of the few people picked to help us understand the travel patterns in this area. Please fill in the questionnaire, tear it off and drop it in the mail. It only takes a minute, and it makes a difference. Thanks!

Detach here and mail

Travel Questionnaire

32 9

Your vehicle (License # 671483) was observed travelling on RT 146 EB, UXBRIDGE
at 6:45AM (Time) on 09/18/91 (WED) (Date).

If vehicle observed is not yours, check here ☐ and return form.

PLEASE PROVIDE THE FOLLOWING INFORMATION:
Your answers are completely confidential.

1. Where were you driving from? (Where did this trip begin?)
   Number ____ Street or nearest intersection ____
   Town/City ____ State ____ Zip Code ____
2. Where were you driving to? (Where did this trip end?)
   Number ____ Street or nearest intersection ____
   Town/City ____ State ____ Zip Code ____
3. How many people including the driver were in the vehicle? ____
4. Which of the following best describes the reason for this trip? (Please check one FROM box and one TO box.)

FROM	TO
1 ☐ Home	1 ☐ Home
2 ☐ Pick up or drop off someone	2 ☐ Pick up or drop off someone
3 ☐ Work	3 ☐ Work
4 ☐ Work-related	4 ☐ Work-related
5 ☐ School	5 ☐ School
6 ☐ Shopping	6 ☐ Shopping
7 ☐ Social activities	7 ☐ Social activities
8 ☐ Recreation	8 ☐ Recreation
9 ☐ Eat out	9 ☐ Eat out
10 ☐ Personal business	10 ☐ Personal business
11 ☐ Other ____	11 ☐ Other ____

5. How long did this trip (one-way) take?
   Approximately ____ minutes
6. Comments ____

After completing, please tear off form at dotted line and drop the survey in the mail. Thank you.

FIGURE 2 External cordon survey travel questionnaire.

Laptop Computer

Recent studies have examined the use of laptop computers to collect traffic data, including license plate data (*7*). Proponents cite the cost savings inherent in direct data entry but recognize that recording rates may not be sufficient for high-volume roads (*8*). During the pilot survey, it was found that the laptop required typing skills beyond those available in the field crew. In addition, the recording rate was lower than for the audiocassette method. So, despite the considerable potential savings in the office processing of the license plate data, the laptop computer method was not used in the main survey.

Audiocassette

The audiocassette method was used successfully in the Boston region in 1977 to conduct the Central Artery Origin and Destination Study (*9*). More than 500,000 license plates were recorded during that survey.

During the pilot survey, it was found that one person taking care and equipped with a good-quality audiocassette recorder and microphone could record at a rate of 1,000 license plates per hour, on the basis of a 15-min test interval. The ability to endure such a high rate varied among field personnel. Both small headways and large gaps in the traffic stream tended to reduce the recording capacity. On the basis of the pilot survey, the audiocassette method was chosen for use at medium-volume locations.

However, problems arose during the main survey. The transcription rate from audiotape was lower than expected, declining to 80 license plates per hour for difficult tapes. In addition, the plate matching rate was significantly lower than for either the notepad or video methods. Therefore, the audiocassette method was replaced after about 2 weeks into the main survey; subsequently, the video method was used at medium- as well as high-volume stations.

It should be noted that a recent study has examined the use of automatic speech recognition for transcribing traffic data from audiocassettes in order to reduce costs (*10*) but without sufficient success to recommend the method at the time of that study.

Video Camera

Closed-circuit television was used as early as 1972 to record license plate data for an O-D survey in Rhode Island (*11*). The use of video cameras to record license plates for this survey was viewed initially with reserve. Previous applications exhibited a labor-intensive and costly process of viewing tapes and transcribing license plates. In addition, the previous applications, although successful, had been relatively site-specific and had involved elaborate and costly setups. For example, the Rhode Island study required extensive equipment assembly and programming, including dedicated platform vehicles, computer equipment, and loop detectors. Similarly, the field equipment used by the California Department of Transportation (Caltrans) to monitor compliance with high-occupancy vehicle (HOV) lane regulations included an instrument van and 110-V source of power (*12*). A more recent Caltrans license plate survey in the Bay Area used expensive off-the-shelf equipment, but the chosen methods required the use of one camera per lane from overhead positions, contributing to fairly high equipment costs (*13*).

In contrast to these applications, the authors were planning to acquire data at a large number of sites under budgetary restrictions that precluded the use of elaborate field equipment. Equipment had to be readily transportable by automobile from one site to another, intended for use by relatively unskilled operators, and inexpensive. It was also critical that the equipment would produce clear, easily readable images of rear license plates on vehicles moving away from the camera at high speeds, under a wide range of ambient light, and at locations where overhead camera positions were not available.

The preliminary training sessions highlighted the difficulty of using unaided visual methods such as notepad, laptop, or audiocassette in reading the rear license plates of vehicles moving away from the observer at high speed on high-volume roads. Therefore, despite reservations regarding the tape viewing and transcription process, the video method was chosen as the only practical way to record plates at the high-volume locations. And, as mentioned previously, the video method supplanted the audio method as the preferred method for medium-volume locations after the main survey was under way.

Plate Transcription and Sampling Methods

After the license data were recorded in the field, the forms (for notepad) and videotapes were collected for transcription. Because of the many plate formats for each state, particular attention was shown toward keying in the plates in the proper formats so that matching with the registry files could be accomplished.

For stations where the notepad and audio methods were used, all plates recorded were key-entered. A subsample of the key-entered plates needed to satisfy sample size requirements for each station was then selected for matching to obtain home addresses for the survey mailout. This was the most efficient method because of the higher proportional sample size at the lower-volume stations where the notepad and audio methods were used.

A different procedure was used at stations where video was used because of the higher cost of transcribing plates from the videotape and the lower proportion of plates needed to satisfy sampling requirements: only the required number of plates needed for matching were transcribed. This was done by first estimating the proportion of traffic required for transcription. For high-volume roads, returned surveys were required from 5 percent of the traffic; therefore, to account for nonmatches and nonresponse, 25 percent of the plates needed to be transcribed. Therefore, the transcriber key-entered all plates in the first 4 min of each 15-min period ($4/15 \approx 27$ percent). This was more efficient than transcribing every fourth plate to achieve the same sample because only 4 min of videotape had to be viewed, rather than the full 15 min. It should be noted that the sampling rate varied by time of day so that equal sample sizes were obtained for three periods: a.m., midday, and p.m.

By the end of the survey, techniques were developed that increased the speed of the video plate transcription process to the point where, when combined with the higher match rate on plates recorded by video, it was viewed as being less costly than transcribing from audiotape. A description of the video equipment used for plate transcription and transcription techniques is included later in this paper.

Plate Matching and Mailing of Survey Forms

Massachusetts, New Hampshire, and Maine plates were matched at CTPS to obtain the home addresses of vehicle owners; the Rhode Island Registry of Motor Vehicles matched the Rhode Island plates. Achieved match rates for video locations approached 85 percent; for notepad locations the match rate averaged 80 percent; for audio stations the match rate averaged about 70 percent.

The matching process appended the home address of the prospective respondent to the license plate files. The files were then transmitted to the mail-merge firm for printing and mailout.

Survey Results

Table 2 gives the number of forms mailed and returned and resultant response rates for each survey station. As shown, more than 110,700 survey forms were mailed; valid responses were received from 29,400 respondents, for an overall response rate of about 26 percent. The response rates varied by station, from a low of 16 percent to a high of 36 percent.

Estimated Survey Costs

The total survey cost, including both consultant and in-house labor and direct expenses, was estimated to be $270,000. Of this amount, direct expenses were approximately $75,000, including $52,000 for printing/postage and $11,000 for video equipment. The overall survey cost estimate includes all labor

TABLE 2 Number of Survey Forms Mailed and Returned

Survey Station	Estimated Sample Size: 1987 ADT	Estimated Sample Size: 1991 Count	Survey Forms: Mailed	Survey Forms: Returned	Response Rate	Proportion of Sample Requirement Met: 1987 ADT	Proportion of Sample Requirement Met: 1991 Count
1	1252	1444	5129	1384	27.0%	111%	96%
2	408	282	1795	364	20.3%	89%	129%
3	255	282	405	76	18.8%	30%	27%
4	368	283	1149	239	20.8%	65%	85%
5	255	245	1127	219	19.4%	86%	89%
6	393	109	690	188	27.2%	48%	173%
7	259	264	1274	394	30.9%	152%	149%
8	797	635	3944	982	24.9%	123%	155%
9	309	289	1720	521	30.3%	169%	181%
10	300	300	1759	417	23.7%	139%	139%
11	1620	1620	7530	1859	24.7%	115%	115%
12	332	263	1486	367	24.7%	111%	140%
13	615	555	3396	735	21.6%	120%	132%
14	299	438	1805	300	16.6%	100%	68%
15	279	232	1613	375	23.2%	134%	162%
16	389	383	1782	449	25.2%	115%	117%
17	1256	1480	6322	1762	27.9%	140%	119%
18	301	425	1506	403	26.8%	134%	95%
19	365	214	1169	379	32.4%	104%	177%
20	871	947	3682	1086	29.5%	125%	115%
21	215	265	941	266	28.3%	124%	100%
22	203	301	1258	375	29.8%	185%	125%
23	255	295	1091	299	27.4%	117%	101%
24	1118	1285	4788	1341	28.0%	120%	104%
25	305	429	1398	440	31.5%	144%	103%
26	729	740	3203	861	26.9%	118%	116%
27	486	455	2019	558	27.6%	115%	123%
28	1166	1185	4417	1125	25.5%	96%	95%
29	332	254	1381	445	32.2%	134%	175%
30	334	297	1057	171	16.2%	51%	58%
31	357	632	2616	632	24.2%	177%	100%
32	211	197	997	247	24.8%	117%	125%
33	243	234	1016	233	22.9%	96%	100%
34	227	237	1214	384	31.6%	169%	162%
35	505	610	2299	600	26.1%	119%	98%
36	316	404	1454	353	24.3%	112%	87%
37	271	374	1353	336	24.8%	124%	90%
38	1458	1543	6266	1448	23.1%	99%	94%
39	486	604	2164	435	20.1%	90%	72%
40	583	778	2655	634	23.9%	109%	82%
41	225	358	1345	389	28.9%	173%	109%
42	279	325	1341	421	31.4%	151%	130%
43	365	442	1610	422	26.2%	116%	96%
44	336	334	1594	435	27.3%	129%	130%
45	871	743	4502	1371	30.5%	157%	184%
46	203	383	907	202	22.3%	100%	53%
47	486	521	2194	683	31.1%	141%	131%
48	101	39	368	134	36.4%	133%	340%
49	648	777	2971	822	27.7%	127%	106%
50	671	725	3259	823	25.3%	123%	113%
TOTAL:	24,900	26,500	113,000	29,400	26.0%	118%	111%

and direct expenses required for survey design and planning; conduct of the pilot survey; all main survey field work, including vehicle occupancy and classification count data collection; survey data transcription, editing, and key entry; and preliminary documentation and analysis. (The estimated total cost does not include the cost of the ATR counts.) On the basis of a total of 29,400 completed surveys, the estimated unit cost was $9.18/completed survey.

CLOSER LOOK AT VIDEO METHODS

This section focuses on the techniques used to collect and transcribe license plate data using video equipment.

Field Equipment

Survey requirements were met by a digital imaging 8-mm camcorder intended for the high end of the consumer market. Although not cheap, at a retail price of approximately $2,300, the camcorder was more than adequate as a camera and an equally suitable playback device. The camcorder fully satisfied the requirements that had been established for its use: it was self-contained and easily portable; it provided clear, stable images of license plates during all daylight hours, even in rain; and it (and ancillary equipment) was readily available at local retail outlets at a cost within the budget.

Site Surveys and Camera Placement and Operation

The general location of each recording site was established by the overall survey design. Within each general location, it was necessary to identify a specific location at which the camera and its operator could be situated safely and at which an appropriate line of sight was provided for viewing traffic. Figure 3 is a copy of a site station map showing recommended camera locations and other features.

At all but one site cameras were set up on or adjacent to the highway shoulder between 10 and 20 ft from the pavement edge. Only traffic on inbound lanes was observed. A single camera was used to record traffic on one lane of two-lane

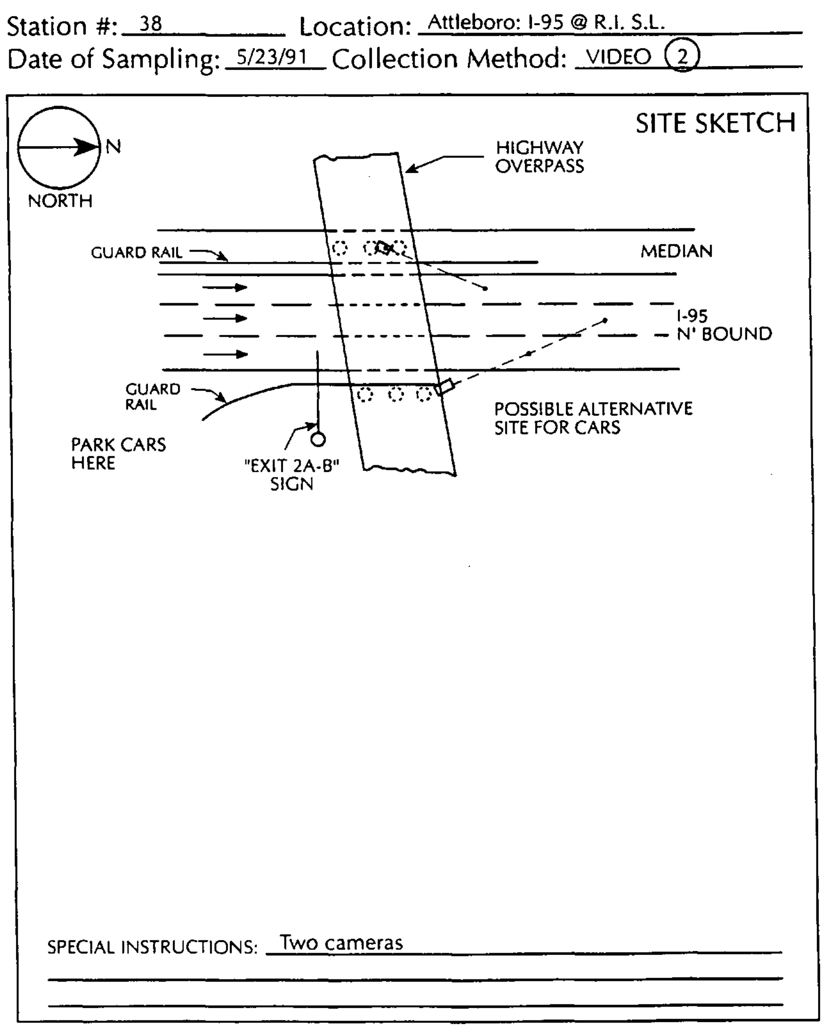

FIGURE 3 Typical external cordon survey station site map.

undivided highways and on both lanes in the inbound direction of four-lane expressways. On six- and eight-lane expressways, two cameras were used: one on the right shoulder and one on the median.

In those instances where only one lane of traffic was to be recorded, the point of sharpest focus was set at the center of the field of view at the center of the observed lane. When a single camera was used to record license plates in two lanes, the point of focus was at the joint of the two lanes.

Office Equipment for Plate Transcription

The camcorder was used initially as the playback device. It was first thought that its high-fidelity playback features would be important in ensuring jitter-free, clear license plate images in the freeze-frame display mode. It turned out that this level of sophistication was unnecessary and that a standard 8-mm player coupled to a remote shuttle controller and an ordinary television provided clear, stable video images of license plates for manual reading and transcription.

Special Techniques To Facilitate Transcription

The techniques just described, such as the use of one camera to videotape two lanes of traffic, and the sampling method significantly reduced the amount of time needed to transcribe license plates from the videotape. An additional technique was developed during the survey to aid the transcription of plates when traffic flow was light.

At high volumes the time intervals between observed vehicles are small, and no problem is experienced in acquiring enough records from a given sampling period. When traffic flows are light, however, the intervals between vehicles or platoons of vehicles are long, and much transcription time can be wasted searching for video frame sequences containing vehicle license plate images.

The most efficient procedure requires the camera operator to block the camera lens for a few seconds at a ½ min or so before the approach of a vehicle or a platoon of vehicles. This allows the viewer to display the tape in the fast-forward/play mode when the blocked (black) interval appears on the screen, shifting into the slow/play mode when images reappear.

Further Refinements

The video camcorder has proved to be a very useful tool for collecting license plate records in moving traffic; indeed, for high-speed, high-volume facilities the video camera appears to be the only feasible means to collect such data. The major disadvantage of the video method is the high cost of manual transcription of license plate data from videotape. The availability of systems that allow the automatic recognition of plate data from videotape will reduce costs substantially. Automation of the transcription process requires the acquisition of high-resolution video images and the availability of computer software capable of locating a license plate against its background and then "reading" the information on the plate.

Such systems recently have been used in Great Britain for O-D studies, but only to record and read European front license plates. Adaptation of these systems for use in this survey was not practicable within the time frame required. However, more recently, one of this paper's coauthors participated in a study to adapt these systems to record American plates. Preliminary results are promising: a 50 percent plate recognition rate was achieved. Further refinements are necessary, but this emerging technology will allow significant reductions in the cost of using video technology in surveys of this kind.

CONCLUSION

Several methods are available to conduct external cordon O-D surveys. For situations in which restrictions regarding traffic delays prevent the conduct of roadside interviews or handing out of postcards, this study has shown that video can be used to obtain the needed information, using off-the-shelf equipment and techniques to minimize costs.

ACKNOWLEDGMENTS

This work was supported by grants from FHWA and FTA, with the assistance and cooperation of the Massachusetts Highway Department.

REFERENCES

1. *Urban Origin-Destination Surveys*. FHWA, U.S. Department of Transportation, July 1975.
2. McKinstry, D. L., and L. G. Nungesser. An Evaluation of On-Site Administered Origin-Destination Survey Methodologies: Postcard Mail-Back Vs. Interview. *Proc., 3rd National Conference on Transportation Planning Methods Applications*, April 1991.
3. Barton-Aschman Associates, Inc. *Phoenix Metropolitan Area External Trip Study, Volume 1: Survey Procedures Manual*. Report FHWA/AZ 86/226-1. Arizona Department of Transportation, Phoenix, Dec. 1986.
4. Stokes, R. W., and T. Chira-Chavala. Design and Implementation of Inter-City Origin-Destination Surveys. In *Transportation Research Record 1236*, TRB, National Research Council, Washington, D.C., 1989, pp. 23–33.
5. Stopher, P. R. Small Sample Home-Interview Travel Surveys: Application and Suggested Modifications. In *Transportation Research Record 886*, TRB, National Research Council, Washington, D.C., 1982, pp. 41–47.
6. Smith, M. Design of Small-Sample Home-Interview Travel Surveys. In *Transportation Research Record 701*, TRB, National Research Council, Washington, D.C., 1979, pp. 29–33.
7. Collura, J., P. McOwen, and D. Demko. *Field Testing of Automated Data Collection in Transportation*. University of Massachusetts, Amherst, Sept. 1984.
8. Bonsall, P. W., F. Ghahri-Saremi, M. R. Tight, and N. W. Marter. The Performance of Handheld Data-Capture Devices in Traffic and Transport Surveys. *Traffic Engineering and Control*, 1988, pp. 10–19.
9. Tippetts-Abbet-McCarthy-Stratton. *Boston Central Artery 1977 Origin-Destination Study*. Executive Office of Transportation and Construction, Commonwealth of Massachusetts, Nov. 1978.
10. Dew, A. M., and P. W. Bonsall. Automatic Speech Recognition—Evaluation of Automatic Transcription of Audio Tape for Registration Plate Surveys. Presented at the 3rd Annual Conference on Survey Methods, Jan. 1990.

11. Danielson, R. F., R. S. Brown, R. M. Hanchett, and W. C. Allinson. A Closed Circuit Television System for Traffic Surveys. *Traffic Quarterly*, Eno Foundation for Transportation, Vol. 28, Jan. 1974, pp. 87–100.
12. Billheimer, J. W., K. Taylor, and C. Slade. *Use of Videotape in HOV Lane Surveillance and Enforcement: Final Report*. Report DOT-T-91-02. Department of Transportation, State of California, March 1990.
13. Fahey, R. M., and B. B. Chuck. *License Plate Videotaping as an Origin and Destination Survey Method Along Highway Corridors*. Public Transportation Branch, District 4, California Department of Transportation, July 1989.

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

Segmentation and Matching of Vehicles in Road Images

MARIE-PIERRE DUBUISSON, ANIL K. JAIN, AND WILLIAM C. TAYLOR

The use of image processing for the segmentation and matching of vehicles in road images is described. Two cameras sense images of the road at different sites in order to estimate the travel time required for a vehicle to travel between two points in a road network. The color images are digitized and analyzed so that moving vehicles can be extracted and segmented into identifiable parts such as the roof, hood, trunk, sides, and wheels. These parts and their attributes such as shape and color are then used to match the vehicles observed from the two different sites. A good match implies that the same vehicle has been sensed by both cameras.

Many research projects are being conducted under the general title of intelligent vehicle-highway systems (IVHS). One of the two primary objectives of IVHS is to reduce travel time by helping the traveler avoid congestion and find the minimum travel time path through a street network. The ability to measure changes in travel time is central to the evaluation of all advanced traveler information systems (ATIS) and advanced traffic management systems (ATMS).

Current techniques for estimating travel time through a network include the floating car and the average speed procedures. These techniques are both labor-intensive and data poor. A vehicle and driver can collect only one observation of the travel time from Point A to Point B in a network on each run. If the network is large or the travel speed low, it can result in only one observation per day in the period of interest (rush hour). Thus, a statistical analysis of changes in the travel time due to a change in traffic signal timing may not be possible because of a limited data set.

License plate matching techniques provide a more robust data base, but they too are labor-intensive. This technique requires observers stationed at each point of interest and coding of the data before a matching algorithm can be exercised. The intent of this research is to develop a procedure that uses remote data collection techniques (using video cameras) and an algorithm that requires no intermediate data coding.

Another potential application of the video image processing technique is the estimation of a real-time origin-destination (O-D) matrix. Several recent studies have looked at algorithms to estimate an O-D matrix using link traffic counts (*1,2*). However, no solution has been found that is accurate but not data-intensive. These algorithms require prior knowledge (or simultaneous data collection) on link volumes and intersection turning volumes. Video image processing has the potential of providing data on the flow from each entry point to all exit points in the network simultaneously—thus an O-D matrix. Although it would require multiple cameras, it would not be labor-intensive for either data collections or data reduction.

This study addresses the data collection problems inherent in current techniques for evaluating ATMS or ATIS. The objective is to explore visual image processing to determine automatically the average travel time from any entry point to any exit point using individual vehicles as probes. The information can then be used in before and after studies to evaluate traffic signal system algorithms and vehicle routing algorithms and to serve as input to driver information systems. To solve this problem, it is necessary to be able to recognize individual vehicles from these images. A block diagram of the system is shown in Figure 1.

Vehicle recognition is a well-studied problem in dynamic scene analysis in the computer vision literature. A segmentation algorithm is applied to the sensed image to extract some features or attributes from the detected vehicle, such as its shape, wheel location, and color. At exit points, images are scanned to search for the same attributes and match them with entering vehicles; a good match means that the same vehicle was seen twice, and the travel time can be estimated by computing the difference in time of the two images.

It is not necessary to match every vehicle, since a large data set can be assembled from the traffic stream, and the present interest is only average travel time data. In this study, the images were obtained where the volume was sufficiently low to have only one vehicle per image. However, the approach will be generalized to the case of multilane highways and higher-volume traffic as the algorithms are refined. The image processing algorithms used here are briefly depicted in Figure 2 and detailed in the rest of the paper.

DETECTION OF MOVING OBJECTS

The simplest method for detecting a moving object in a fixed background is by image subtraction. In a method proposed by Jain et al., the detection of a moving object requires the object to have moved by an amount at least equal to its length in the image (*3*). The boundary of the object can then be extracted perfectly. This assumption was considered too restrictive, since it requires a large field of view that could contain two vehicles. To extract the details of individual vehicles, it would be desirable to have the vehicle occupy as much of the image as possible. In other words, a compromise between the perfect contour of the moving object and a small field of view is required. The method developed in this study does not produce a perfect contour of the moving object, but

Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Mich. 48824.

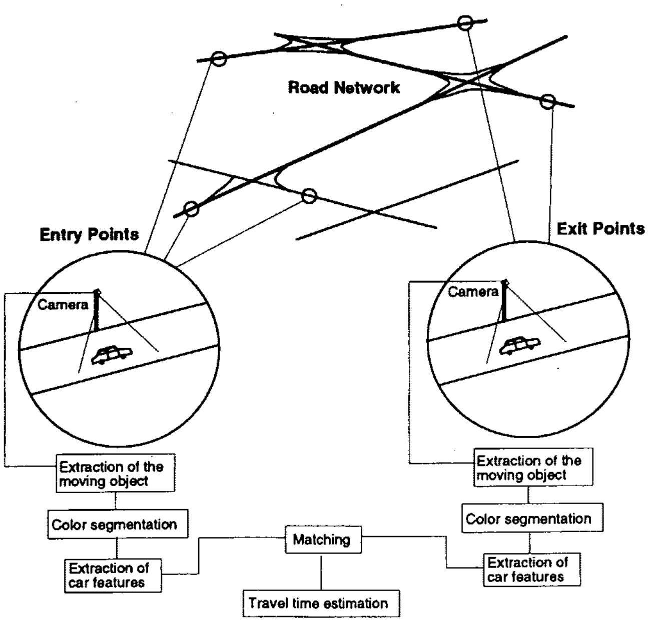

FIGURE 1 Road network and associated sensors and processing modules.

it has the advantage of using a smaller field of view, resulting in a larger area for the vehicle in the image frame.

Two color images are input to the algorithm, the second image being a few frames later than the first image in the sequence (interframe time is 1/30 sec). The images are 512 × 512, and each pixel is represented by a vector of intensity values in the three-dimensional color space (red, green, and blue). Each color component is quantized into 256 levels. Let

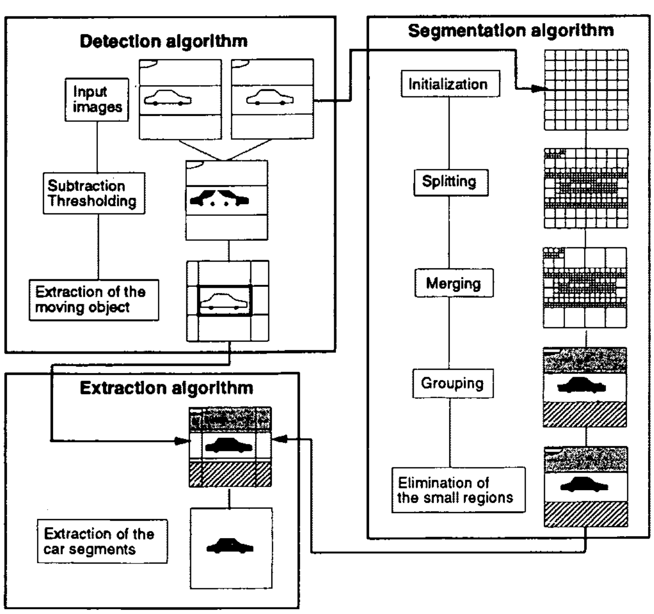

FIGURE 2 Image processing algorithms.

$\mathbf{f}^1$ and $\mathbf{f}^2$ represent the two images captured at times t_1 and t_2:

$$\mathbf{f}^1 = [\mathbf{f}^1(i,j)] = [f_R^1(i,j), f_G^1(i,j), f_B^1(i,j)] \tag{1}$$

$$\mathbf{f}^2 = [\mathbf{f}^2(i,j)] = [f_R^2(i,j), f_G^2(i,j), f_B^2(i,j)] \tag{2}$$

where $\mathbf{f}^1(i,j)$ is the intensity vector at pixel (i,j). The two images are subtracted at each pixel and thresholded to produce a binary image I in the following manner.

If

$$|f_R^1(i,j) - f_R^2(i,j)| \geq t \vee |f_G^1(i,j) - f_G^2(i,j)| \geq t \vee |f_B^1(i,j) - f_B^2(i,j)| \geq t \tag{3}$$

then

$$I(i,j) = 1, \qquad \text{else } I(i,j) = 0 \tag{4}$$

where $\vee$ refers to the logical OR operator. The parameter t is called the threshold; this case, t has been fixed at 30.

Thresholding the difference image produces a binary image in which most of the "on" pixels (pixels with the value 1) belong to the moving vehicle. Some other pixels take the value 1 because of the noise or the movement in the background (for example, wind in the tree foliage or the grass), but these pixels either are too isolated or belong to small groups that cannot be part of a vehicle. A rectangular area containing the moving vehicle is extracted by summing over the rows and columns of the binary image $I(i,j)$ and finding the rows and columns where these sums are significant. The row sums must be greater than 1/10 of the image length, and the column sums must be greater than 1/20 of the image height. These numbers would be modified if a larger field of view is used where the vehicle sizes appear to be smaller in the image.

Figure 3 summarizes the process of extracting moving objects. Figures 3*a* and 3*b* show the red components of frame i and frame $i + 3$ in the input image sequence. Figure 3*c* shows the binary image obtained after subtraction and thresholding and Figure 3*d* shows the rectangular area surrounding the moving vehicle in frame $i + 3$.

The initial vehicle image data base used in this study contains 22 road images. Each color image contains one vehicle. These images were captured by recording a road scene outside the engineering building on the Michigan State University campus. A videocassette recorder was used in combination with an image digitizer on a workstation to acquire the still images. For each of the 22 images, a bounding rectangle for the moving object was correctly identified. Thus, this method can be used to accurately determine which part of the image contains the moving vehicle.

COLOR SEGMENTATION

Color segmentation is the process that divides the image into homogeneous regions (called segments) using the color information at each pixel. Three methods are commonly used to perform this task.

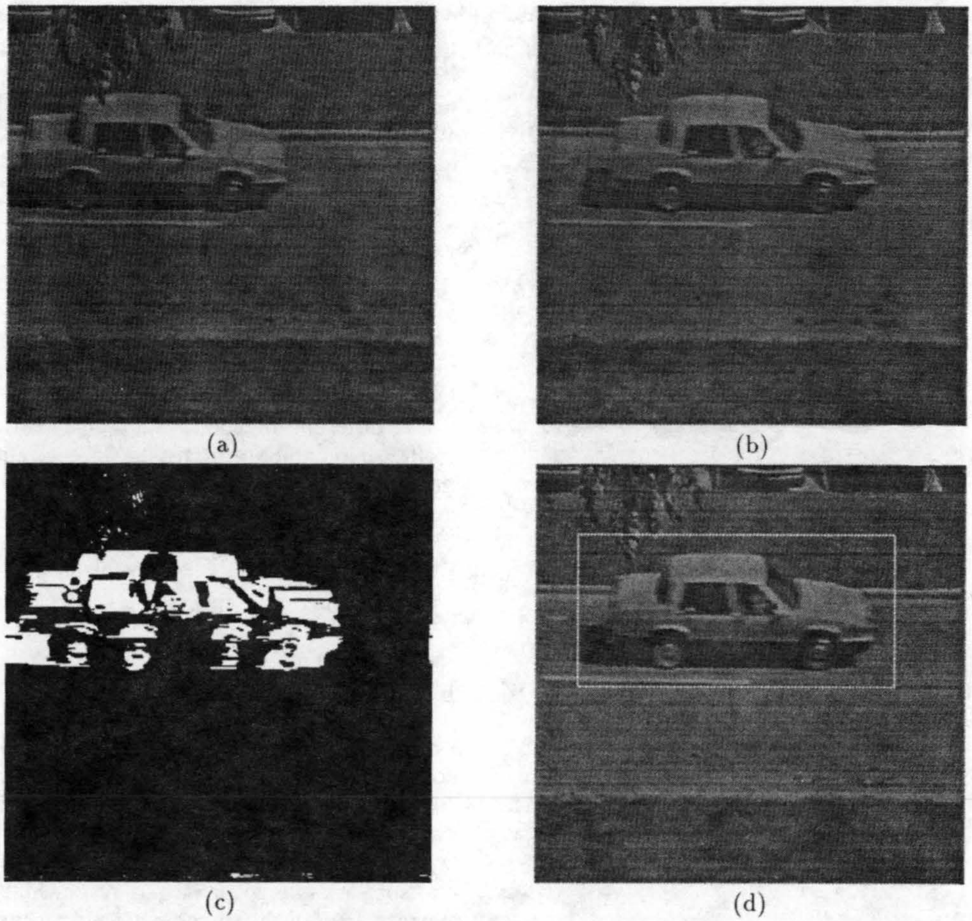

FIGURE 3 Detection of moving vehicle: *a*, Frame *i*; *b*, Frame $i+3$; *c*, output of subtraction and thresholding; and *d*, rectangular area outlining moving object.

Clustering

Clustering consists of finding clusters of points in the three-dimensional color space and assigning each cluster to a different segment (*4*). The Euclidean distance metric is commonly used to find points that are sufficiently similar to each other to define a cluster. The main disadvantage of this method is that the number of clusters (or segments) needs to be specified in advance.

Recursive Region Splitting

Recursive region splitting takes a region of the image and determines a threshold in one feature by finding the best peak in the set of feature histograms (initially the entire image is considered) (*5*). The subregions defined by this threshold are then further segmented if necessary. Ohlander used the following features: R, G, B (original tristimuli red, green, and blue), I, H, S (intensity, hue, and saturation), and the linear combination Y, I, Q (color system for TV signal). See the work by Ballard and Brown for more details on these color components (*6*). This algorithm can be very unstable because of the nonlinearity of the functions used to compute the hue and saturation components. Ohta et al. showed, using a Karhunen-Loeve transformation, that the transformed features

$$I1 = (R + B + G)/3$$
$$I2 = (R - B)/2$$
$$I3 = (2G - R - B)/4 \qquad (5)$$

produced the best segmentation results (*7*).

Split-and-Merge Algorithm

The main advantages of the split-and-merge algorithm (*8*) are that the number of segments need not be known in advance, allowing the background to be arbitrarily complicated, and the detected homogeneous regions are connected—the spatial location of a pixel plays a role in its assignment to a region. In this project, a modification of this algorithm was used as described in the following. The color segmentation algorithm is divided into five steps.

Initialization

The $N \times N$ image (where $N = 2^k$) is divided into a number of square subimages. For example, a 512×512 image is divided into 64 squares that are 8 × 8. In each square subimage, the mean and variance of the intensity values for each of the three color components are computed as follows:

$$m_R = \frac{1}{z^2} \sum_{i=x}^{x+z} \sum_{j=y}^{y+z} R(i,j)$$

$$m_G = \frac{1}{z^2} \sum_{i=x}^{x+z} \sum_{j=y}^{y+z} G(i,j)$$

$$m_B = \frac{1}{z^2} \sum_{i=x}^{x+z} \sum_{j=y}^{y+z} B(i,j) \qquad (6)$$

and

$$v_R = \frac{1}{z^2} \sum_{i=x}^{x+z} \sum_{j=y}^{y+z} [R(i,j) - m_R]^2$$

$$v_G = \frac{1}{z^2} \sum_{i=x}^{x+z} \sum_{j=y}^{y+z} [G(i,j) - m_G]^2$$

$$v_B = \frac{1}{z^2} \sum_{i=x}^{x+z} \sum_{j=y}^{y+z} [B(i,j) - m_B]^2 \qquad (7)$$

where (x,y) are the coordinates of the upper left corner of the square and z is the length of the subimage (initially $z = 8$).

Splitting

Subimages that are not homogeneous are further split into four squares of length $z/2$ and new means and variances are computed for each new square. This is repeated until the individual subimages have reached an area of 4 pixels. A subimage is not homogeneous if

$$(\sigma_R^2 + \sigma_G^2 + \sigma_B^2) > t_\sigma \qquad (8)$$

Merging

If four adjacent subimages have not been split in the second step and satisfy the homogeneity condition

$$\{|m_{R_i} - m_{R_j}| + |m_{G_i} - m_{G_j}| + |m_{B_i} - m_{B_j}|\} < t_m \quad \forall i,j \qquad (9)$$

then they are merged to form a larger subimage of length $2z$.

Grouping Subimages To Form Regions

Grouping subimages eliminates the arbitrary region boundaries imposed by the initialization process. Square subimages are merged into regions if they satisfy the criterion

$$\{|m_R - M_R| + |m_G - M_G| + |m_B - M_B|\} < t_g \qquad (10)$$

where the M's are the means for the region currently being formed. A region grows by the addition of neighboring subimages when they satisfy this criterion. The region means are updated every time a new subimage is added to the region. When no more subimages can be added, a new region is created.

Deleting Small Regions

During this segmentation process, some very small regions are created, which are then removed by merging them with neighboring larger regions. This is done on the basis of a minimum size threshold, t_s.

The different threshold values given earlier are parameters to the segmentation program. Since no universal segmentation algorithm has been found yet, these parameters must be adjusted to properly segment the specific images being analyzed. In this study, the parameters were fixed at $t_m = 50$, $t_v = 50$, $t_g = 40$, and $t_s = 250$ for all the images used in the experiments.

Figure 4 shows the segmentation results for the vehicle image shown in Figure 3*b*. A total of 38 segments have been identified in this image. One segment corresponds to the road and two others correspond to the grass. The curb and marking on the road can also be extracted. Different segments appear for the leaves in the tree and the vehicles parked in the parking lot in the background. Meaningful segments are also extracted from the vehicle. The roof, hood, trunk, two windows, and wheels can be distinguished. The use of color information in the split-and-merge process is responsible for this excellent performance of the segmentation algorithm.

EXTRACTION OF VEHICLE

Once the image has been segmented and the moving object has been detected, those segments that are part of the moving object can be extracted from the segmented image to obtain a segmented image of the vehicle alone. This is done by considering all the segments and discarding those that are outside the bounds of the rectangular area. The segments for which

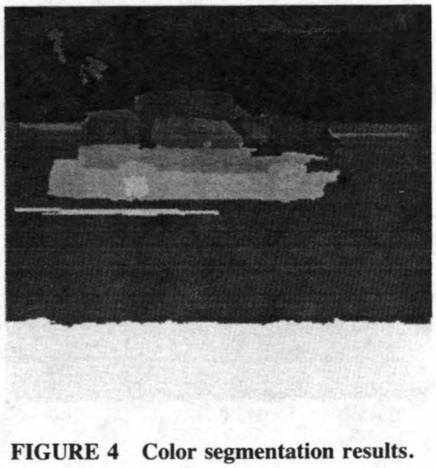

FIGURE 4 Color segmentation results.

the intersection of their area with the bounding rectangular area is less than 80 percent are also eliminated. Also, elongated regions that belong to either the curb or the markings on the road can be suppressed by looking at each column of the image and setting to the background color the regions (in the background) of height fewer than 15 pixels. The result of this process is shown in Figure 5. Note that the vehicle in Figure 5 was extracted by taking the intersection of the bounding rectangle in Figure 3*d* and each of the 38 segments in Figure 4. It can now be seen that only the segments belonging to the vehicle are retained.

This algorithm was tested on 22 car road images, each containing one vehicle. The vehicle segments have been extracted successfully in all the images on which this algorithm has been tried. Results for two other vehicles are shown in Figures 6 and 7. Again, only regions belonging to the vehicle are kept, and these regions represent meaningful parts of the vehicle.

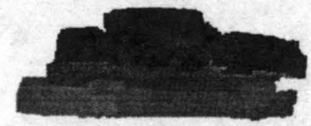

FIGURE 5 Extraction of vehicle from background.

INITIAL MATCHING RESULTS

For each vehicle, a low-level description now exists in terms of a number of segments. These segments can be used by a matching algorithm to determine if a given vehicle has been sensed at an earlier time. The matching algorithm used is based on a graph matching approach (*6*). A graph is built to obtain a high-level description of each vehicle. A node in this graph corresponds to a segment in the segmented image. The node attributes are the number of pixels in the segment, the average color of the segment (three components: red, green, and blue), and the perimeter of the segment. Two nodes are connected by an edge in the graph if the corresponding two segments are neighbors in the segmented image. The arcs in the graph also have attributes, which are the distance between the centroids of the two corresponding segments and the orientation of the line linking those two centroids. An example of this relational graph is shown in Figure 8.

Matching two vehicles is equivalent to matching the corresponding two relational graphs. A similarity measure can be computed by comparing the node attributes and the arc attributes in the two graphs. This is done as follows: the best matching two nodes are found by comparing their attributes, and the next best matching two nodes are found by comparing their attributes and their relationship (link attributes) with the other nodes already matched. This process stops when all the regions in at least one of the vehicles have been matched.

Among the 22 road images that were available, 3 showed the same vehicle at different times (Vehicles 5, 15, and 21). For experimentation purposes, a known vehicle was driven three times around the building. A similarity measure was obtained for each possible pair of vehicles. Table 1 presents the similarity measures for some of those pairs of segmented images. The similarity measures are 0.95 between car5 and car15, 0.95 between car5 and car21, and 0.94 between car15 and car21. The high similarity measures indicate that the vehicles must be the same, and indeed they are, as explained earlier. The other similarity measures, however, are quite low, indicating that the two vehicles are not the same.

One of the remaining questions is the definition of the level of the similarity measure required to conclude that a match has been made. In Table 1, car8 has similarity measures of 0.90 with car5 and car21 and 0.86 with car15, yet these are not matches. Using a threshold of 0.91 would eliminate these false matches, but if the threshold were set at 0.95, the true match between car15 and car21 would have been missed.

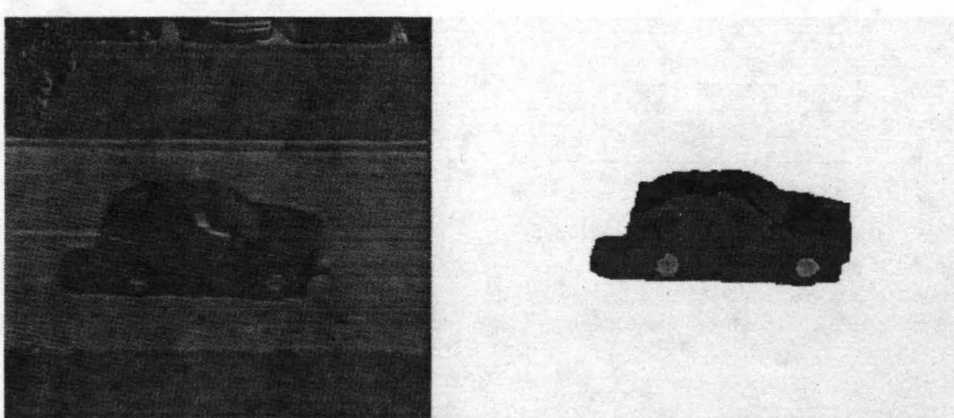

FIGURE 6 Results of region segmentation, Vehicle 1: *left*, original image; *right*, region segmentation of vehicle only.

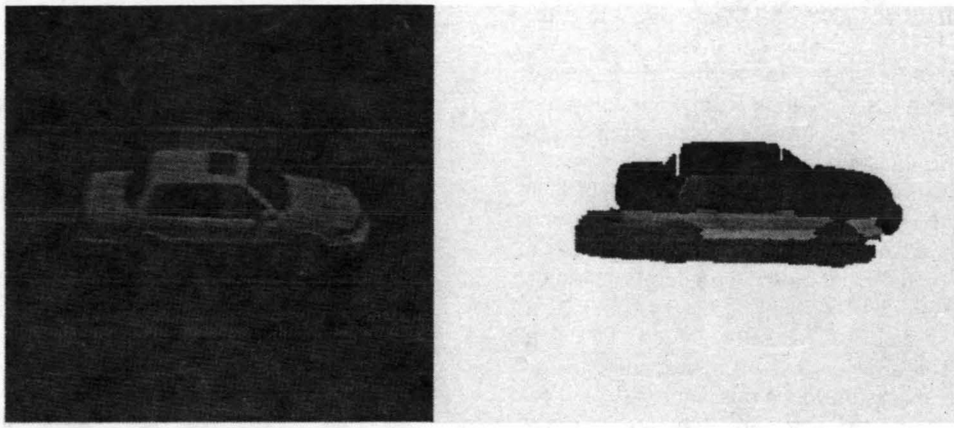

FIGURE 7 Results of region segmentation, Vehicle 2: *left*, original image, *right*, region segmentation of vehicle only.

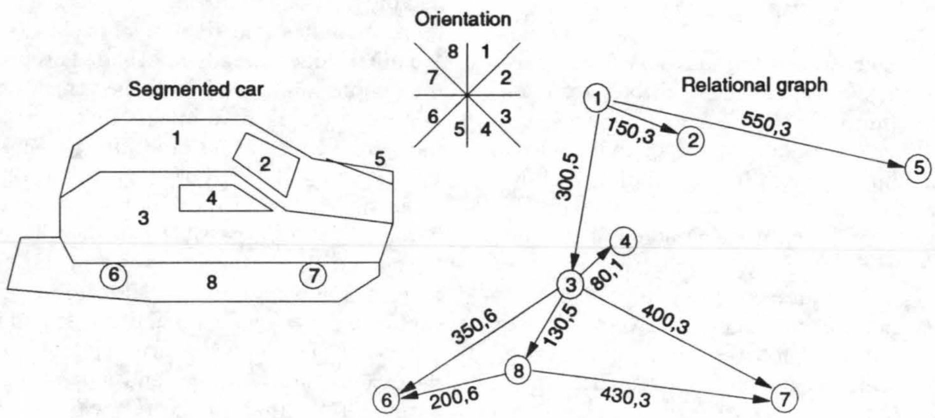

FIGURE 8 Relational graph.

TABLE 1 Initial Matching Results

	car4	car5	car6	car7	car8	car13	car15	car17	car18	car21
car4	1.00	0.08	0.34	0.42	0.09	0.39	0.07	0.12	0.07	0.05
car5		1.00	0.04	0.36	0.90	0.06	**0.95**	0.58	0.34	**0.95**
car6			1.00	0.06	0.31	0.12	0.05	0.32	0.17	0.05
car7				1.00	0.09	0.27	0.07	0.17	0.48	0.08
car8					1.00	0.54	0.86	0.58	0.05	0.90
car13						1.00	0.05	0.10	0.08	0.26
car15							1.00	0.56	0.31	**0.94**
car17								1.00	0.42	0.56
car18									1.00	0.05
car21										1.00

These results show that the segmented image of a vehicle can be used for matching. Once pairs of vehicles have been matched, the travel time for a given vehicle can be computed if the time it was sensed by the first camera and the time it was sensed by the second camera are known.

CONCLUSIONS

It is possible to extract the moving vehicle from the background and segment it into meaningful regions such as the roof, hood, trunk, and wheels. The quality of the segmented images is very good, because color has been used during the segmentation process. These segmented regions—along with their shape, color, and the relationships between them—describe the structure of the vehicle fairly well. A relational graph is then constructed to summarize this information. A graph matching algorithm is used to determine whether a vehicle observed by the second camera at an exit point has been seen by the first camera at an entry point of the road network.

A number of matching algorithms have been developed and presented in the literature (9–11). Future work will involve an investigation of a better matching algorithm for this problem and extending the process to handle images of multilane highways. The problem of handling multilane highways can be solved either by ignoring the vehicles that do not appear in the front lane (they would still need to be separated from the vehicle in the first lane in the image, which is a difficult image segmentation problem) or by monitoring each lane separately, which would require many more cameras. It is almost certain now that the first solution will be chosen.

ACKNOWLEDGMENTS

This research was supported by a grant from the U.S. Department of Transportation through the Great Lakes Center for Truck Transportation Research.

REFERENCES

1. Moshe, B.-A. Methods To Combine Different Data Sources and Estimate Origin-Destination Matrices. In *Transportation and Traffic Theory* (N.H. Gartner and N. H. M. Wilson, eds.), 1987, pp. 459–481.
2. Cremer, M., and H. Keller. A New Class of Dynamic Methods for the Identification of Origin-Destination Flows. *Transportation Research*, Vol. 21B, No. 2, 1987, pp. 117–132.
3. Jain, R., and H.-H. Nagel. On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 1, No. 2, April 1979, pp. 206–214.
4. Jain, A. K., and R. C. Dubes. *Algorithms for Clustering Data.* Prentice-Hall, Englewood Cliffs, N.J., 1988.
5. Ohlander, R., K. Price, and R. Reddy. Picture Segmentation Using a Recursive Region Splitting Method. *Computer Graphics and Image Processing*, Vol. 8, 1978, pp. 313–333.
6. Ballard, D. H., and C. M. Brown. *Computer Vision.* Prentice-Hall, Englewood Cliffs, N.J., 1982.
7. Ohta, Y.-I, T. Kanade, and T. Sakai. Color Information for Region Segmentation. *Computer Graphics and Image Processing*, Vol. 13, 1980, pp. 222–241.
8. Horowitz, S., and T. Pavlidis. Picture Segmentation by a Tree Traversal Algorithm. *Journal of the Association for Computing Machinery*, Vol. 23, No. 2, April 1976, pp. 368–388.
9. Price, K., and R. Reddy. Matching Segments of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 1, No. 1, Jan. 1979, pp. 110–116.
10. Davis, L. Shape Matching Using Relaxation Techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 1, No. 1, Jan. 1979, pp. 60–72.
11. Cheng, J. K. *Algorithms for Matching Relational Structures and Their Applications to Image Processing. Ph.D. thesis. Purdue University, West Lafayette, Ind., Dec. 1980.*

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

Nonresponse Bias and Trip Generation Models

PIYUSHIMITA THAKURIAH, ASHISH SEN, SIIM SÖÖT, AND EDWARD CHRISTOPHER

There is serious concern over the fact that travel surveys often overrepresent smaller households with higher incomes and better education levels and, in general, that nonresponse is nonrandom. However, when the data are used to build linear models, such as trip generation models, and the model is correctly specified, estimates of parameters are unbiased regardless of the nature of the respondents, and the issues of how response rates and nonresponse bias are ameliorated. The more important task then is the complete specification of the model, without leaving out variables that have some effect on the variable to be predicted. The theoretical basis for this reasoning is given along with an example of how bias may be assessed in estimates of trip generation model parameters. Some of the methods used are quite standard, but the manner in which these and other more nonstandard methods have been systematically put together to assess bias in estimates shows that careful model building, not concern over bias in the data, becomes the key issue in developing trip generation and other models.

Nonresponse bias is a well-recognized problem in sample surveys. In the field of transportation, where much of the planning effort rests on estimates of parameters obtained from models, it becomes pertinent to discuss the manner and the extent to which nonresponse bias affects the quality of estimates.

There are two types of nonresponse:

1. *Total nonresponse* occurs if some individuals or households simply do not respond to the survey. Then bias could occur if the preferences, values, or behavior of the nonrespondents are different from those of the respondents on whom estimates are based.
2. *Item nonresponse* occurs if parts of the survey instrument are not completed. A particularly unpleasant form of this in travel surveys occurs when respondents forget to record all trips. The result is that some individual households appear to have taken fewer trips than they actually did.

Trip generation models, whether they are cross-classification models or continuous models, are typically linear models, and estimates can be viewed as least-squares estimates. Since this is not entirely obvious for cross-classification models, a proof is given in a later section. It is well known that if certain conditions are satisfied, least-squares estimates are unbiased. On the other hand, if the model does not satisfy these conditions, bias will occur even if there is a 100 percent response rate. These conditions are satisfied by the model if the functional form of the model is correct and all important explanatory variables are included.

Because categorical models do not have any problems with their functional form, and weighting and related issues are taken care of, the authors prefer categorical trip generation models. This preference is discussed in a later section. Therefore, the issue that remains when assessing bias in estimates from categorical trip generation models is whether the model includes all the relevant independent variables or at least all important predictors.

Methods of assessing bias caused by omitting important variables are demonstrated on a trip generation model of trip circuits constructed from data that were collected by a survey in Lake County, Illinois, in the northern part of the Chicago metropolitan area. Two types of approaches are available for this:

1. Try out different variables in addition to those included in the model. Since the variables chosen to include in the model were fairly standard, it should be expected that such analyses have been carried out by model builders in the past.
2. Use various diagnostics for correct model specification. The kind of diagnostic methods used and the results are shown later.

Yet another diagnostic tool is available in the case of categorical models. Since the dependent variable is counted, it is reasonable to expect that its values are approximately Poisson. This assumption is made explicitly or implicitly in nearly all contingency table (discrete multivariate) literature in statistics—and cross-classification models are special cases of such models. An examination of the empirical distribution of the number of circuits revealed an unexpected phenomenon. It showed that in one-member households the distributions were very close to Poisson, whereas in larger households the distributions were different, and the difference appeared to indicate the presence of item nonresponse.

A variable that has been used as a surrogate for the differences between respondents and nonrespondents is also examined. This analysis does not lead to any results that could contradict the conclusions that the estimates obtained from the model are unbiased.

The overall conclusion, with respect to the Lake County model that the authors built and examined, is that the model shows no signs of substantial bias due to variable omission

P. Thakuriah, A. Sen, and S. Sööt, Urban Transportation Center, M/C 357, University of Illinois at Chicago, Reliable Building Annex, Suite 700 South, 1033 West Van Buren Street, Chicago, Ill. 60680. E. Christopher, Chicago Area Transportation Study, 300 West Adams Street, Chicago, Ill. 60606.

and that total nonresponse has no noticeable adverse effect. Yet item nonresponse remains a serious problem. This leads to a natural suggestion that individual household members, rather than households, should be asked about trip-making behavior. Clearly this is easier with mail-out/mail-back surveys than with telephone or perhaps even home interview surveys.

Logit as well as gravity models are generalized linear models (*1*). However, it should be noted that in most estimation procedures for the gravity model, the dependent variable is effectively the number of trips going from one zone to another. The effect of total nonresponse at the household level is akin to item nonresponse at the zonal level. Thus the arguments in this paper do not apply to such gravity model parameter estimation procedures. However, a discussion paralleling the one in this paper for trip generation models can be made for logit models.

SOME THEORETICAL CONSIDERATIONS

It is known that when linear least squares is used to estimate regression coefficients (and certain conditions are met), randomness in the independent variable values is not a necessity for unbiasedness. This fact has been exploited to combat bias in sample surveys (*2*). However, because of its importance for this argument and to make the paper more or less self-contained, this fact is demonstrated in following paragraphs.

Cross-classification trip generation analysis is shown to yield least-squares estimates that are the same for a wide range of realistic weights. This leads to the reasons that the authors prefer cross-classification models for trip generation modeling.

Condition for Unbiasedness

A linear regression model can be written as

$$y = X\beta + \boldsymbol{\varepsilon} \tag{1}$$

Equation 1 implies that the linear relationship between the variables in the X-matrix (of independent variables) and y (number of trips) is the same for all households except for minor fluctuations (given by the error term, $\boldsymbol{\varepsilon}$). In the sequel weighted least squares shall be considered for its greater generality and also because trip generation models are concerned with counted data, which typically render ordinary least squares inappropriate.

The weighted least-squares estimate, $\boldsymbol{b}$, of β, is

$$\boldsymbol{b} = (X'WX)^{-1}X'Wy \tag{2}$$

where $w_1, \ldots, w_n$ are positive weights and W is the diagonal matrix whose diagonal elements are $w_1, \ldots, w_n$ [i.e., $W = \text{diag}(w_1, \ldots, w_n)$]. If $\boldsymbol{b}$ is to be an unbiased estimate of β, then one of a set of three conditions, collectively called the Gauss-Markov conditions, must be met by Equation 1. This condition is

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{o} \tag{3}$$

where $\boldsymbol{\varepsilon}$ equals $(\varepsilon_1, \ldots, \varepsilon_n)'$ and n is the number of households. From Equation 2,

$$\begin{aligned} \boldsymbol{b} &= (X'WX)^{-1}X'W(X\beta + \boldsymbol{\varepsilon}) \\ &= (X'WX)^{-1}X'WX\beta + (X'WX)^{-1}X'W\boldsymbol{\varepsilon} \\ &= \beta + (X'WX)^{-1}X'W\boldsymbol{\varepsilon} \end{aligned}$$

Therefore, when

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{o} \tag{4}$$

it follows that $\mathrm{E}(\boldsymbol{b}) = \beta$, showing that $\boldsymbol{b}$ is unbiased. Note that each component ε_i of $\boldsymbol{\varepsilon}$ relates only to the ith observation and not to all observations. Thus, if each observation is believed to be valid and the regression model chosen is valid, $\mathrm{E}(\varepsilon_i) = 0$. Then, of course Equation 4 would hold.

One way in which Equation 4 will fail to hold is if an explanatory variable is omitted from the analysis. The condition will also be violated if the algebraic form of the model is incorrect or does not apply to all the observations. No other condition is required for the unbiasedness of regression coefficients (*3,4*). In particular, the values of independent variables do not need to be random. The problem then is to develop the model without leaving out any important explanatory variable and by specifying the correct form of the model.

Categorical and Continuous Trip Generation Models

As mentioned earlier, two kinds of trip generation models are in common use. One kind is the so-called categorical or cross-classification model. An example of such a model is a table in which the rows correspond to, say, several levels of household size and the columns correspond to different levels of income. The entries themselves are average number of trips (or trip circuits) made by households in that category (the trip rates). Such models are equivalent to regression models with dummy variables but without an intercept term, as will be shown. In the categorical model presented later, household size and number of workers per household are called factors that have several levels. The other kind of trip generation model is in the form of a single equation containing mainly continuous independent variables. They will be called continuous regression models.

Typical categorical trip generation models are also linear regression models. There are two ways in which this can be seen:

- Approach 1:

Let y_i be the number of trips taken by the ith household, and let

$$x_{ij} = \begin{cases} 1 & \text{if } i\text{th household is in category } j \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Here category means a cell of the cross-classification table. Then a regression using y_i's as dependent values and x_{ij}'s as independent values and using no intercept gives exactly the table entries.

• Approach 2:
Alternatively, let y_i be the number of trips in category j and

$$x_{ij} = \begin{cases} \text{number of households in category } i & \text{if } i = j \\ 0 \quad \text{otherwise} \end{cases} \tag{6}$$

The only difference between Approaches 1 and 2 is that Approach 1 is a disaggregate version of Approach 2. In the first approach, there is one y_i for each i, whereas in the latter approach, y_i is a summation over all i's in category j.

To show that least squares would give exactly the same estimates as cross-classification would, consider Approach 1. Let $\boldsymbol{x}_j$ be the jth column of the matrix X of elements x_{ij}. Each such column contains only 0's and 1's, the 1's in column $\boldsymbol{x}_j$ being only in those rows for which the observation belongs to the jth category. Note that there is no intercept term.

Since the number of trips taken in each household is obtained by counting, it is also called a counted variable. When the dependent variable is counted, it typically has a Poisson distribution and weighted regression is usually called for to avoid a violation of the second of the Gauss-Markov condition given by

$$E(\varepsilon_i^2) = \sigma^2 \tag{7}$$

where $i = 1, 2, \ldots, n$ and σ is a positive constant. For a categorical model, these weights w_i and ω_j would be constant for each category j. Indeed, the condition of constancy of weights within each category is about as general a condition as the authors could require. The authors assume this condition in the sequel.

If $j \neq \ell$, $\boldsymbol{x}_j$ will contain a 1 only in those rows where $\boldsymbol{x}_\ell$ has a 0, since a household belongs to one category. Therefore $\boldsymbol{x}_j' W \boldsymbol{x}_\ell = 0$. The number of 1's in each $\boldsymbol{x}_j$ is the number n_j of households in category j. Therefore, $\boldsymbol{x}_j' W$ is a vector consisting of n_j nonzero terms, each of which is ω_j and $\boldsymbol{x}_j' W \boldsymbol{x}_j = n_j\omega_j$. The matrix $X'WX$ is, therefore, diagonal, diag $(n_1\omega_1, \ldots, n_k\omega_k)$ with diagonal elements $n_j\omega_k$. Consequently,

$$(X'WX)^{-1} = \text{diag}\,(n_1^{-1}\omega_1^{-1}, \ldots, n_k^{-1}\omega_k^{-1}) \tag{8}$$

If y_i is the number of trips taken by the ith household, $\boldsymbol{x}_j W y$ is the total number, t_j, of trips taken by household i in category j, times ω_j. Consequently, $X'Wy$ is a vector whose jth element is $t_j\omega_j$. Hence, $\boldsymbol{b} = (X'WX)^{-1}X'Wy$ is a vector, the jth element of which is $\omega_j t_j/\omega_j n_j = t_j/n_j$, the trip rate for the jth category. Thus least squares using Model 5 gives exactly the same estimates one gets from categorical analysis—regardless of weighting so long as the weights are constant within a category. Approach 2 can be handled in a similar way.

When adequate data are available, the authors prefer categorical trip generation models over continuous models for the following reasons:

• No explicit weighting is required for categorical models.
• Nonlinearity of the model is usually not an issue for such models.

The only major shortcoming of categorical models is that, since each factor has several levels, the number of independent variables is large and, consequently, large numbers of observations are required. However, this is not a problem in transportation studies. Thus, categorical models would usually be more reliable.

EXAMPLE

A categorical trip generation model was developed. A brief description of the data is given, the model is presented, and the adequacy of this model to give unbiased trip generation estimates is checked.

The data used in the project were obtained by the Chicago Area Transportation Study in October 1989 from a mail-out/mail-back survey of Lake County in northeastern Illinois (Chicago Area Transportation Study, unpublished data, Aug. 1990). Lake County is a rapidly growing region of approximately half a million residents in north suburban Chicago, bordering Wisconsin, Lake Michigan, and Cook County. The county is a low-density suburban area encompassing large estates as well as sizable low-income areas.

The data set includes information on two types of variable: transportation-related variables, including the origin and destination of every trip, its purpose, travel time, mode used, automobile occupancy (for automobile trips) and walking distance if transit modes are involved, and census-related variables, such as number of persons per household, age, vehicle availability, gender, employment status, occupation, and income.

Respondents to the Lake County household travel behavior survey reported their household travel information for one of two days. The first travel date was October 12, 1989. Subsequently, reminder letters were mailed to those households that had not yet returned filled-in questionnaires and whose mail had not been returned by the post office as undeliverable. Two substitutes for the first travel day that had passed (Thursday, October 19 or October 26, 1989) were suggested to these households. A total of 9,143 questionnaires were sent out and 2,480 households returned usable questionnaires (a return rate of 27.1 percent).

TRIP GENERATION MODEL

A categorical household trip generation model was developed on the basis of two commonly used household socioeconomic variables, household size and number of employees per household. These variables are essentially traditional, having withstood the test of numerous studies. However, since such trip generation models have existed for a while and presumably are extensively investigated, it is unlikely that an important variable has not been considered at some stage. The purpose of this paper is not to indicate which socioeconomic variables in the trip generation model are more useful predictors of household travel but to examine such a model in the light of the earlier discussion to see if the estimates are biased.

To illustrate the earlier discussion, the cross-classification trip generation model has been presented both in cross-tabular form (in Table 1) and in the form of a regression model with no intercept and with dummy variables (in Equation 9, as in

Approach 1). In this trip generation model, the base household trip generation rates have been defined in terms of trip circuits. Trip circuits may be defined as the round-trip movements by household members that begin and end at home. Trip circuits, rather than trips, were the focus of this study because the trip generation models of the Chicago Area Transportation Study are in terms of trip circuits, and some consistency with those models was seen to be desirable. Trip generation calculations were obtained after developing cross tabulations of households stratified by occupants and employees. (Trip circuits have also been referred to in the literature as primary trips and trip chains.) The model obtained is

$$\begin{aligned}\text{Trip circuits}_i = {} & 0.81x_{i(0,1)} + 1.90x_{i(0,2)} + 2.08x_{i(0,3)} + 3.00x_{i(0,4)} \\ & + 2.00x_{i(0,5)} + 1.04x_{i(1,1)} + 1.89x_{i(1,2)} + 2.29x_{i(1,3)} \\ & + 2.61x_{i(1,4)} + 3.26x_{i(1,5)} + 2.33x_{i(1,6)} + 2.11x_{i(2,2)} \\ & + 2.45x_{i(2,3)} + 2.85x_{i(2,4)} + 3.84x_{i(2,5)} + 4.19x_{i(2,6)} \\ & + 3.24x_{i(3,3)} + 4.12x_{i(3,4)} + 4.38x_{i(3,5)} + 5.12x_{i(3,6)} \\ & + 4.47x_{i(4,4)} + 5.75x_{i(4,5)} + 6.83x_{i(4,6)} + 6.67x_{i(5,5)} \\ & + \varepsilon_i \qquad R^2 = .669,\ s = 3.12 \end{aligned} \tag{9}$$

where the dependent variable is the trip circuit rates in the ith household and the subscript within parenthesis for each x indicates the number of occupants and the number of workers in the ith household [for instance, $x_{i(1,0)}$ is the dummy variable that takes the value 1 if the ith household has one member who is not employed, and 0 otherwise].

The estimate b_j of β_j for each of the $j = 1, 2, \ldots, J$ categories in the trip generation model (Equation 9), then, is exactly equal to the mean trip circuit rate of that category (as in Table 1). In Table 1, there are 36 categories, whereas in Equation 9, there are 24 independent variables. Therefore, Equation 9 is estimating the trip rate in only some of the categories. The independent variables that were deleted had columns of only 0's in the X matrix (there were no households in these cells), and the columns of the matrix were, consequently, linearly related. Omitting these independent variables to avoid singularity also had the intuitive reasoning that there is no sense in estimating trip rates where there are no household entries at all.

TABLE 1 Categorical Trip Generation Model

Workers	Summary Statistic	Household Size 1	2	3	4	5	6
0	Mean:	0.81	1.90	2.08	3.00	2.00	—
	Var:	0.75	2.45	2.63	9.00	12.00	
	N:	128	189	12	3	3	
1	Mean:	1.04	1.89	2.30	2.61	3.26	2.33
	Var:	0.67	2.29	3.44	3.01	3.76	6.52
	N:	248	228	107	123	61	15
2	Mean:	—	2.11	2.45	2.85	3.84	4.19
	Var:		2.01	3.40	4.52	4.41	12.16
	N:		461	217	240	66	16
3	Mean:	—	—	3.24	4.12	4.38	5.12
	Var:			3.61	8.66	6.17	8.93
	N:			99	56	26	12
4	Mean:	—	—	—	4.47	5.75	6.83
	Var:				9.03	11.11	0.80
	N:				32	12	5
5	Mean:	—	—	—	—	6.67	—
	Var:					15.47	
	N:					6	

The '—' indicates that there are no entries in that category.

Statistical Assessment of Bias

The issue in this section is how one decides if substantial bias is present in the estimates. As mentioned, there are two ways in which bias can occur:

1. If a nonlinear situation is represented with a linear model, or
2. If an important variable is left out of the model.

Because the present model is categorical, where the estimated number of trips in the category is the mean number of trips in each category, no difficulty occurs because of Reason 1.

To detect whether a variable has been left out is always difficult. The statistical literature suggests two ways to check for omitted variables:

1. Examine outliers (which often show the need for additional variables), and
2. Examine if there is a relationship between the predicteds and the residuals of the model (because the orthogonality of the predicteds and the residuals require the unbiasedness of model estimates).

To look for outliers, the authors examined plots of residuals and the Studentized residuals against the predicteds, which indicated a few data points as possible outliers. These points had Studentized residuals of 2.78, 4.06, and 4.58 and DFFITS of 1.96, 1.01, and 1.02. [See works by Sen and Srivastava (*3*) and Belsley et al. (*4*) for a definition of Studentized residuals and DFFITS, the latter being a statistic commonly used to identify influential points.] The cutoff for DFFITS of 0.20, which is a consequence of typically suggested formulae, was not used because it drew attention to 5 percent of the 2,480 data points. Each of these three outliers represented households with large number of household members and no trip circuits. There were no working members in the first two and only one worker in the third. The outliers appear to be due to "natural causes." We found no compelling reason to suspect that the model in Equation 9 did not include all important variables.

As mentioned, it is also useful to see if there is a relationship between the residuals and the predicteds because the deletion of an important variable can cause the residuals of the model to have nonzero expectations. The plots did not reveal any systematic relationship. To corroborate this further, a second regression model was developed in which the dependent variable was the predicteds ($\hat{y}_i$) and the independent variable was the residuals (e_i) from Equation 9. The R^2 of .0003 indicated close to no fit between the predicteds and the residual. But, as pointed out elsewhere (*3*), the lack of a relationship between $\hat{y}_i$ and e_i does not conclusively lead to the decision that the model is unbiased, because patterns between residuals and the predicteds are not always apparent.

Poisson and Item Nonresponse

Another check for bias that is available for categorical models also did not suggest bias, but it did yield unexpected results. Variables whose values are obtained by counting something are known to have a Poisson distribution if the items that are counted are statistically independent. Although trips are not exactly independent, it is usually conjectured that the Poisson distribution approximately holds. Moreover, this assumption is consistent with the customary assumption made in traffic engineering that flows on links are Poisson—something that has been observed to be approximately true. A Poisson assumption underlies nearly all contingency table (discrete multivariate) analyses in statistics.

From Table 1 it can be seen that the mean-to-variance ratio deviates from 1 without any clearly discernible pattern, although the deviation is more marked for categories with a small sample size. To get a clearer idea of why the deviation from Poisson was occurring, the theoretical Poisson probabilities were computed for each category with means from the trip generation model (Equation 9) as the Poisson parameter.

The comparison of the theoretical Poisson distribution for each category with the actual data points revealed the possibility of item nonresponse (see Table 2; the first set of numbers for each category is the theoretical distribution for that category, the second set is the empirical distribution). In households with one member, no matter whether that member is a worker or a nonworker, the distribution of actual trip circuits conforms approximately to that of the theoretical Poisson distribution of trip circuits and in fact underestimates the theoretical frequency of no trip circuits. As household size increased, the number of households reporting zero trip circuits was far greater than what the theoretical distribution for that category predicted. This finding indicates an important way in which the data collected by sample surveys may bias trip generation information—that when one person in the household fills out the survey questionnaire, he or she might miss recording trips made by other members in the household.

The statistical techniques used reveal that although the data from which the trip generation model was developed showed some item nonresponse, there is no reason to suspect that the estimates from the model are biased (in the sense of total nonresponse). The authors, therefore, do not believe that any

TABLE 2 Comparison of Theoretical Poisson and Empirical Distributions of Trip Circuits by Category

Household	Workers	Number of Trip Circuits										
Size		0	1	2	3	4	5	6	7	8	9	10+
1	0	56.80	46.15	18.75	5.08	1.03	–	–	–	–	–	–
		54	50	20	2	2	–	–	–	–	–	–
1	1	100.86	104.42	54.05	18.66	4.83	1.00	–	–	–	–	–
		77	130	69	7	0	1	–	–	–	–	–
2	0	28.28	53.72	51.03	32.30	15.34	5.83	1.85	–	–	–	–
		53	21	54	27	22	11	–	–	–	–	–
2	1	34.43	65.09	61.52	38.77	18.32	6.93	2.18	0.59	–	–	–
		54	40	63	34	27	5	5	–	–	–	–
2	2	55.73	117.76	124.40	87.61	46.27	19.55	6.89	2.08	0.55	0.13	0.11
		89	33	173	95	52	16	2	0	0	0	1
3	0	1.49	3.11	3.24	2.25	1.17	0.49	–	–	–	–	–
		2	3	3	1	2	1	–	–	–	–	–
3	1	7.20	19.43	26.21	23.59	15.92	8.60	3.87	1.49	0. 50	–	–
		25	12	22	25	10	7	3	2	1	–	–
3	2	18.69	45.84	56.18	45.91	28.14	13.79	5.63	1.97	0.60	0.16	–
		48	14	52	43	34	15	7	1	2	1	–
3	3	3.87	12.54	20.33	21.98	17.81	11.55	6.24	2.89	1.17	0.42	–
		16	2	4	34	21	13	6	1	1	1	–
4	1	9.05	23.61	30.81	26.80	17.49	9.13	1.48	–	–	–	
		20	12	24	32	20	7	6	2	–	–	–
4	2	13.94	39.68	56.45	53.55	38.10	21.68	10.28	4.18	1.49	0.47	–
		51	4	51	48	46	16	13	2	2	2	–
4	3	0.91	3.73	7.70	10.59	10.92	9.01	6.19	3.65	1.89	0.86	0.45
		12	1	3	7	7	9	3	5	5	3	1
4	4	0.34	1.64	3.66	5.46	6.09	5.44	4.06	2.59	1.45	0.72	–
		8	0	0	3	2	4	4	7	3	1	–
5	1	2.34	7.62	12.43	13.52	11.03	7.19	3.91	1.82	0.74	–	–
		9	0	9	16	15	4	4	3	1	–	–
5	2	2.02	7.05	12.29	14.27	12.43	8.67	5.03	2.51	1.09	0.42	–
		10	2	5	13	19	6	7	2	1	1	–
6	1	1.45	3.39	3.95	3.09	1.78	0.84	0.33	0.11	0.03	0.008	–
		8	0	2	3	4	5	0	0	0	9	–
6	2	0.24	1.02	2.13	2.97	3.11	2.61	1.82	1.09	0.57	0.26	0.19
		4	0	1	2	5	4	1	2	0	0	1

important predictor of trip rates is missing from the model. However, work on an additional variable is described in the next section.

Analysis of Late Responses

The variable that was explicitly examined is response itself, if it is assumed that late respondents represent a point on the continuum between respondents and nonrespondents. Under this assumption, those independent variables that would be important in distinguishing between early and late respondents to the survey would also be useful predictors of the difference between respondents and nonrespondents.

The approach presented in this section is based on the premise that individuals who respond late to a survey, and from whom a response is elicited only after a reminder letter had been sent or a follow-up telephone call had been made, are "closer" to nonrespondents than to respondents because of the prodding needed to get the response from them. This is a fairly standard assumption. In a landmark study by Filion using this technique (*6*), nonrespondents were considered as persons who resist the initial waves of the questionnaire. All respondents form a continuum from highly motivated to unmotivated individuals. Each wave probes deeper into the core of the nonrespondents, and the continuum is indicative of the direction and extent of total nonresponse bias. Consequently, Filion claims that extrapolation over successive waves will reflect the characteristics of the hard core of the nonrespondents (*6*) [see also work by Armstrong and Overton (*7*) and Finn et al. (*8*)].

The data gathering design by Chicago Area Transportation Study led most naturally to this part of the analysis. Households that responded to the Lake County household travel behavior survey were divided into two groups, early and late, on the basis of the date on which they reported their travel information. Households that filled the survey form using October 12 or an earlier Thursday (October 5, 1989) as the travel day, were termed as the early respondents. Households that responded using Thursdays after October 12 were termed late respondents. The number of early respondents according to this definition was 1,907 households, and the number of late respondents was 573 households.

The new independent variable introduced to serve this purpose was a dummy variable z_j for each j, which takes the value 1 if the ith household in j responded early to the survey and 0 if the household responded late. A single dummy could have been used for each of the households for which trip circuits are predicted. However, a dummy variable for each j allows greater flexibility. The regression needs to be weighted, but earlier results—that estimates from cross-classification models are weighted least-squares estimates for a wide class of weights—are true only for estimates themselves, not for their standard errors. Thus, for testing purposes specific weights need to be specified.

For simplicity, the authors decided to proceed in a different way. Since the estimates of regression coefficients are trip rates, it was decided to simply compare trip rates using as variance estimates the usual sample variances. That is, a standard t-test was used to compare trip rates for early and late respondents in each category. The results are given in Table 3. A theoretical justification of this approach would follow the lines given by Sen and Srivastava (*3*).

Only one of the differences in Table 3 was significant at a 5 percent level (although in that case the significance was at a substantially lower level). Given 21 separate t-tests, getting one significance at a 5 percent level is about what should be expected if there is no relationship. Although the early re-

TABLE 3 Trip Circuit Rates (TCR) for Early and Late Respondents for Complete Model

Workers	Respondent Type	Household Size 1	2	3	4	5	6
0	Early TCR:	0.83	1.97	1.75			
	Late TCR:	0.73	1.51	2.75			
	Obsns.	[102, 22]	[158, 31]	[8, 4]	[2, 1]	[2, 1]	—
	t value:	0.59	1.59	-0.98			
1	Early TCR:	1.06	1.89	2.40	2.68	3.28	2.22
	Late TCR:	0.95	1.88	1.92	2.40	3.22	2.50
	Obsns.	[228, 56]	[180, 48]	[83, 24]	[91, 32]	[43, 18]	[9, 6]
	t value:	0.29	0.08	1.16	0.78	0.10	-0.21
2	Early TCR:		2.19	2.65	2.93	3.67	4.46
	Late TCR:		1.90	1.89	2.51	3.05	3.00
	Obsns:	—	[331, 130]	[160, 57]	[193, 47]	[46, 20]	[13,3]
	t value:	—	1.86	3.12*	1.06	0.99	0.73
3	Early TCR:			3.25	4.56	4.59	
	Late TCR:			3.22	2.36	4.00	
	Obsns.	—	—	[76, 23]	[45, 11]	[17, 9]	[11, 1]
	t value:			0.06	1.98	0.56	
4	Early TCR:				4.54	6.50	7.00
	Late TCR:				4.17	4.25	8.00
	Obsns:	—	—	—	[26, 6]	[8, 4]	[3, 2]
	t value:				0.25	1.32	-1.73
5	Early TCR:					7.50	
	Late TCR:					6.25	
	Obsns:	—	—	—	—	[2, 4]	
	t value:					0.37	

spondents, in most cells, indicated higher trip rates than late respondents, a reasonable conclusion still is that this variable is not too important.

An examination of the empirical distribution of trip circuits made by early and late respondents in each category also showed that the two groups essentially had the same pattern in the household trip circuits. This corroborates the conclusion that (assuming late respondents are closer to nonrespondents than to early respondents the model is reasonably unbiased and that there is no significant difference between the respondents and the nonrespondents.

CONCLUSION

Estimates from categorical trip generation models are unbiased if certain conditions are met by the model. The focus of developing trip generation models, therefore, shifts from being concerned about overrepresentation or underrepresentation in the data of households possessing certain socioeconomic characteristics to checking the model to see if these conditions are met. A cluster of checks is suggested to verify whether the model gives unbiased estimates of household trip generation parameters.

An empirical analysis was done to illustrate how this bias assessment may be done in practice. A series of statistical diagnostic tools was used, including outlier analysis and examination of the residuals and the predicteds from the model. These tools indicated that the model does not have substantial total nonresponse bias. However, the empirical distribution of only one-member households closely followed the theoretical distribution for such categories of households. This indicated that there is possible item nonresponse in households in which one person records trip information for the other household members. Finally, on the premise that late respondents are closer to nonrespondents than early respondents on a continuum of respondents to nonrespondents, an analysis was done to check if the two groups of respondents are sufficiently different. The results agreed with the earlier analysis that the model is reasonably free of total nonresponse bias.

REFERENCES

1. McCullagh, P., and J. A. Nelder. *Generalized Linear Models*. Chapman and Hall, London, England, 1989.
2. Särndal, C. E., and T. K. Hui. Estimation for Nonresponse Situations: To What Extent Must We Rely on Models? In *Current Topics in Survey Sampling* (D. Krewski, R. Platek, and J. N. K. Rao, eds.), Academic Press, San Diego, Calif., 1981.
3. Sen, A., and M. Srivastava. *Regression Analysis: Theory, Methods and Applications*. Springer-Verlag, New York, 1990.
4. Belsley, D. A., E. Kuh, and R. E. Welsch. *Regression Diagnostics—Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, Inc. New York, 1980.
5. Rao, C. R. *Linear Statistical Inference and its Applications*. John Wiley & Sons, Inc., New York, 1973.
6. Filion, F. L. Estimating Bias Due to Non-response in Mail Surveys. *Public Opinion Quarterly*, Winter, 1975–1976, pp. 482–492.
7. Armstrong, J. S., and T. S. Overton. Estimating Non-response Bias in Mail Surveys. *Journal of Marketing Research*, Vol. 14, Aug. 1977, pp. 396–402.
8. Finn, D. W., C. Wang, and C. W. Lamb. An Examination of the Effects of Sample Composition Bias in a Mail Survey. *Journal of Marketing Research Society*, Vol. 25, No. 4, Oct. 1983, pp. 331–338.

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

Use of Direct Data Entry for Travel Surveys

JERRY C. N. NG AND PAUL M. SARJEANT

Telephone interviews are a popular way to collect survey information. In the traditional paper-and-pencil method, the interviewer records the responses on paper forms, and later the information is input into a computer file for error checking and further analysis. The quality of the data depends not only on interviewing skill but also on the ability of the interviewer to write legibly, the accuracy of the data entry staff, and the feedback process in reporting ambiguous or incorrect information. Data of higher quality can be obtained by having the interviewer directly enter the data into a computer file as the interview proceeds. Direct data entry would minimize data entry errors while enhancing quality control and the overall processing of the data. The design and use of a direct data entry system in the conduct of a major household travel survey in the greater Toronto area are discussed. Besides obtaining good-quality data, the DDE software also improved sample control, the rate at which interviews were completed, and the monitoring of interviewers' performance and progress. To identify the costs and benefits of direct data entry, extensive comparisons are made with a survey that was similar in terms of survey area and questionnaire design but conducted using the paper-and-pencil method.

In the fall of 1991, a comprehensive household travel survey—the 1991 Transportation Tomorrow Survey (TTS)—was conducted in the greater Toronto area. The 1991 survey was intended to update a survey performed in 1986, which was also called the Transportation Tomorrow Survey. The earlier survey has been the primary data source for transportation planning in the greater Toronto area. Both the 1986 and 1991 surveys were conducted by telephone and preceded by an advance letter informing households about the survey. The survey area includes the regional municipalities of Durham, York, Peel, Halton, and Hamilton-Wentworth and the municipality of metropolitan Toronto (Figure 1). Households were sampled randomly using telephone listings supplied by Teledirect, a subsidiary of Bell Canada. The most significant difference in the conduct of the two surveys was in the area of automation, especially in the use of a direct data entry (DDE) system in the 1991 survey.

The 1986 TTS used the traditional paper-and-pencil technique [despite some promising results from a prototype DDE system that was tested as a part of the pilot study for that survey (*1*)]. The interviewer recorded the responses on standard paper forms. After the interview supervisor checked them visually, the interview forms were passed to a data entry team for data entry and then on to a coding team. The sampling of households was done manually, and no logic or range checks were performed on the data collected until the whole household interview had been fully coded.

Because of the size of the survey (about 61,700 completed household interviews were performed), the data entry and coding processes were unable to keep pace with the interviewing, a problem that worsened as the survey progressed. In some cases, households with ambiguous information were not called back until weeks after the original interview. Although the 1986 survey data were reliable and informative, it took more than a year of postprocessing and checking before the data were available for detailed analysis. At the conclusion of the 1986 survey, the need to improve both quality control and the efficiency of the data processing was clearly identified, so the use of DDE for the 1991 survey was recommended (*2*).

WHY A TELEPHONE SURVEY?

There were a number of reasons for the original decision to conduct a telephone survey instead of a self-reporting mail-back survey. The experience in a number of Canadian urban travel surveys had been that telephone surveys, with sufficient interviewer training, can achieve a relatively high response rate. Telephone interviewers can clarify confusing questions when interviewees register confusion or resistance, which tends to increase the response rate. In mail-back surveys, the respondent burden is high, especially on detailed travel survey questionnaires. Therefore, mail-back surveys can underrepresent groups that are not fully literate in English or not accustomed to filling out complex forms; examples of such groups are seniors and recent immigrants.

It was understood that telephone surveying tends to underreport discretionary and non-home-based travel. To assess this, the 1986 TTS was followed up with a mail-back diary survey to a random sample of the interviewed households. The results of this survey indicated that the telephone survey collected peak-period trips just as well as the mail-back trip diary survey (*3*). In addition, measures such as proper follow-up calls to those persons whose behaviors could not be reported accurately by the person who answered the telephone can (and do) result in improved overall trip reporting.

BENEFITS OF DIRECT DATA ENTRY

The benefits achieved by the DDE system, as implemented for the 1991 TTS, were twofold. First, it enhanced the quality of the data collected during the course of the survey. The

J. C. N. Ng, Data Management Group, Joint Program in Transportation, University of Toronto, 42 St. George Street, Toronto, Ontario M5S 2E4 Canada. P. M. Sarjeant, BA Consulting Group Ltd., 111 Avenue Road, Suite 604, Toronto, Ontario M5R 3J8 Canada.

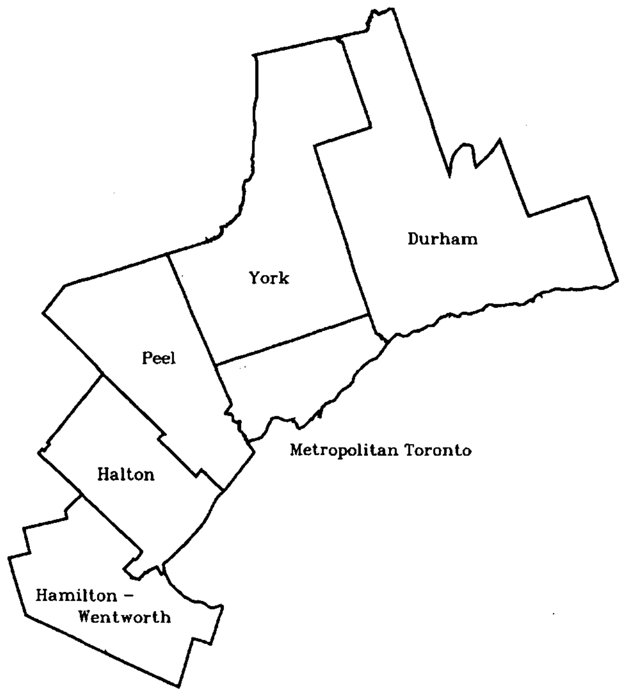

FIGURE 1 Survey area: greater Toronto area.

ability of the interviewer to enter the information as the interview proceeded eliminated the need for data entry after the interview had been completed. This capability saved time and effort and minimized the potential for data entry errors. Having the interview in electronic format also enhanced the efficiency of sample and quality control. The DDE system was fully integrated with a sample control software system. The interview status (for example, "successfully completed," "call back required," or "answering machine encountered") of each household was monitored daily by the sample control software. By integrating the sample control and interviewing software, a household could be scheduled to be interviewed on a specific date at a specific time if necessary.

Second, it dramatically reduced the amount of time required for the postprocessing of the survey data. The postprocessing task for the 1986 survey included data entry, coding of transit routes, and assigning X and Y geocodes to all location information. For the 1991 survey, DDE eliminated the need for post-survey data entry. By implementing look-up tables of transit routes (and their interconnections) within the DDE software, transit routes were verified and coded as the interview proceeded. Geocoding of location data (e.g., trip origin, destination, transit transfer point) was also aided through the use of look-up tables of street and municipality names. Logic and range checks were also part of the DDE system. Fundamental errors, such as driving under age or without a driver's license, were noted and corrected immediately, thus minimizing the number of call backs and post-survey edits required. The substantial reduction in elapsed time to code, check, and validate the collected data was viewed as one of the greatest benefits achieved through the use of this software (*4*).

SOFTWARE DESIGN

The DDE system has been written using a DOS-based data base management software package to run on stand-alone IBM-compatible personal computers (PCs). The key design criterion was to automate as much as possible of the data collection and verification process, without sacrificing speed to the point that the interview would become unacceptably long. The speed of the DDE system is evidenced by the fact that it can be run satisfactorily on a 16Mhz 80286-based microcomputer. Key features of the system include an integrated household sample selection and control capability, a built-in interview script, and cross referencing of information. The data verification features include on-line logic and range checking and look-up tables for verification of key items, including school names, transit routes, municipalities, and street names.

Sample Selection and Control

At the beginning of every interview session, each interviewer receives a sample group of households on a floppy diskette. The sample is generated by the sample control software running on a central sample control and data processing machine. The sample issued to each interviewer consists of households with three interview status levels, scheduled times to call back, prior unsuccessful attempts to make contact (e.g., no answer), and no prior contact histories. The sample file is first copied from the floppy diskette to the hard disk of the PC being used by the interviewer, and subsequent work is performed directly from the hard disk. To avoid loss of information during an interview, information is saved onto the hard disk at the end of each data field entry, and the complete file is copied back onto the floppy diskette at the end of each interview.

The DDE software automatically sequences the households for the interviewer by first drawing households with scheduled call-back times, followed by the physical order of the households in the sample file. There are two types of scheduled call backs: those set by the interviewer and those set automatically by the DDE software. An interviewer can schedule a call back if a household requests to be interviewed at a later date and time or if additional information is to be provided at a later time. Automatic call backs are set by the DDE software if an interviewer encounters a busy signal, no answer, or an answering machine. In the latter case, the interviewer is also given a brief message to leave on the answering machine. Up to eight attempts are made to contact each household.

An interviewer can also specify which household to interview or review. There are two ways to do this: interviewers can select one of the interviews they have worked on earlier during the same interview session, or a supervisor logged into the system (i.e., running the software in supervisor mode) can specify a household by its telephone number. Running the software in supervisor mode is most useful in dealing with a household that calls in response to a message left on its answering machine and to perform post-interview corrections and edits.

Working Screens

Every working screen in the DDE system is divided into three sections. Relevant information collected during the interview is constantly updated and displayed in the top portion of the screen. For example, on the trip data screen, the person's name, age, gender, and home and work addresses are shown while trip information is collected. The middle portion of the screen is the active area where the interviewer keys in the required information. Full screen editing capability is available in this middle section. The bottom section of the screen contains instructions to the interviewer, including the interview script for the current data item, available options, and valid response codes. Function keys are used to jump from one screen to another for quick editing and review. Most responses are subject to logic and range checking. When the keyed data are in error, a warning message is displayed and the interviewer is instructed to either change the keyed entry or confirm that it is the intended response.

The five main working screens are described in the following.

Household Selection Screen

The Household selection screen is the first screen presented to the interviewer after logging into the system. Interviewers may choose to either interview the next household or review a previously interviewed household. Supervisors may choose two additional options: selecting a specific non-English-speaking household or selecting a household by its telephone number.

After a choice is made, the DDE software locates an appropriate household in the interview sequence or the household specified by the user. The household's family name, address, and telephone number are then displayed, along with information about any previous contact attempts. This contact history includes the time and date of previous calls, the reasons for having to call back (such as no answer, line busy, answering machine encountered, or incomplete interview), and any memo messages left behind by previous interviewers. These memo messages can be reviewed or edited at any time during the interview. The objective is to familiarize the interviewer with the household as much as possible before contact is made.

Using a series of messages, the screen guides the user through the process of contacting members of the appropriate household, finding out if they received the presurvey mailing informing them of the purpose of the survey, and determining if they are willing to participate at this time. If so, the interview proceeds to the household data screen.

Household Data Screen

Household information collected during the survey includes confirmation of the household's address, the dwelling type, the number of household members, and the number of vehicles available for the use of household members. Because of the multicultural nature of the greater Toronto area, the language used for the interview is also indicated by the interviewer on this screen (Figure 2). If the household indicated a preference to be interviewed in a language other than English, the interviewer would terminate the interview, indicating to the software that a call back was desired by someone with skills in that language.

The information collected on this screen, particularly the number of people, sets logical and sequence conditions for the collection of person and trip information.

Person Data Screen

Personal information is collected for one person at a time, for each member of the household, before the travel information is collected. Person data include the person's name or some other identification label (such as mother, father, or respondent), age, gender, possession of a driver's license, employment and student status, and the person's usual place

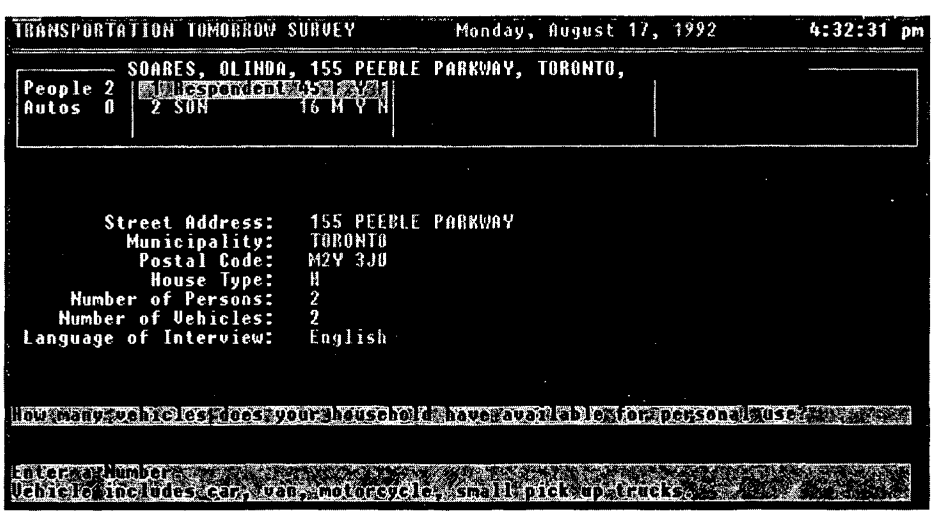

FIGURE 2 Household data screen.

of work or school, if applicable (see Figure 3 for example of screen). The work or school address is entered as a street address, an intersection, or a monument name. Monument names are identifiers for particular buildings, landmarks, or attractions. This information is verified as it is entered using the look-up tables described in the following. If the person is employed, the interviewer asks whether free parking is available at work. Employment, student, and trip information are not collected for persons younger than 11.

Trip Data Screen

The aim of the DDE system is to reconstruct sequentially the complete 1-day travel activities of each person in the household. To do this the interviewer establishes whether each person made any trips on the travel day. This travel day is usually the weekday previous to the day on which the survey takes place. If an individual made one or more trips, the information collected is the first origin of the day (usually the home) followed by all subsequent trip destinations.

The trip end point is entered as a street address, intersection, or monument. This information is again verified using the look-up tables. Because a person's daily trip activities are often made up of his or her home, work, and school locations, trip destinations can also be specified with a home, usual place of work, or usual place of school choice. The address information collected in the household and person data screens are automatically transferred to the trip records. This feature not only speeds up the interviewing process, it also eliminates duplication of effort during geocoding and gets rid of a potential source of error.

Trip purpose, start time, and mode of travel are also recorded. The start time is checked to ensure that trips are being recorded sequentially. If the mode of travel is automobile driver, the software checks for the possession of a driver's license and the availability of a private vehicle.

The trip data screen (Figure 4) contains the most information of all the working screens. Not only does it display the current person's personal and trip information, it also shows the same trip information for all other members of the household. If several members of the household traveled to the same location, a simple copy command can duplicate an entire trip. Full screen editing capability is available so that any trip record can be edited. However, only trips for the current active person can be inserted or deleted.

Transit Data Screen

One of the requirements of the survey was to collect detailed routing information for transit trips. Thus, a special screen was designed solely for these trips (Figure 5).

The transit data screen is activated when public transit is specified as the mode of travel. Besides individual transit routes, access and egress modes to and from transit are also recorded. When the first or last route is commuter rail or subway and the access or egress mode is automobile driver, automobile passenger, or taxi, the transfer station name is also recorded. On-line transit coding is performed using look-up tables of transit routes and station names. Transit routes are identified by either their formal or alias names or their route numbers. Every route is also checked for connectivity with the previous route.

Because a commuter will often use the same routes to return to their initial origin, an option to reverse the routings of a previous transit trip is made available to the interviewer.

LOOK-UP TABLES

Providing on-line detailed tables of helpful information on the different working screens is one key to the smooth operation of the DDE system. It reduces key strokes while enhancing data quality by minimizing spelling errors and ensuring that a description complete enough to allow accurate geocoding has been recorded. For example, when recording a street address, it is important that its street type (e.g., "Av-

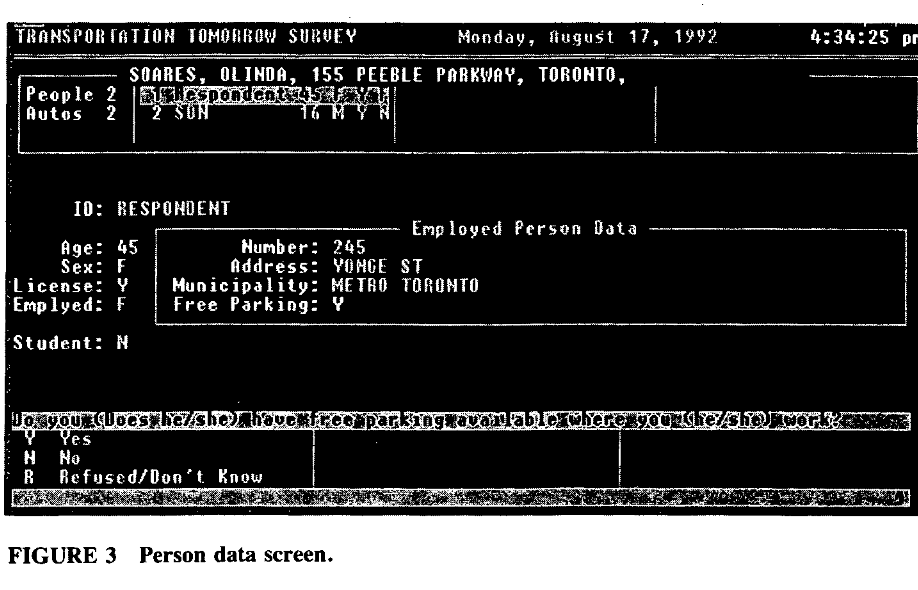

FIGURE 3 Person data screen.

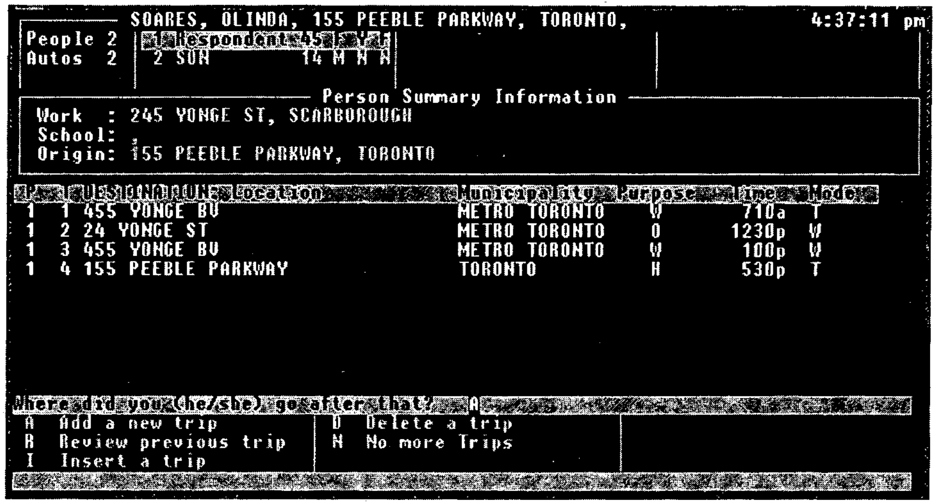

FIGURE 4 Trip data screen.

enue," "Road," "Court") and direction (e.g., "East," "West") are also recorded. Because many streets have similar or exact names ("King Street" being the worst offender in the survey area), it is also important to have the interviewer confirm the correct municipality for the street address. Another example is in the recording of schools and community colleges for which the campus name and the institution type are vital.

The tables are presented in pop-up "browse and select" windows. There are five of these windows for listings of municipalities, streets, schools, transit routes, and transit stations. All of these windows can be operated in two ways. First, as the interviewer enters information character by character, the software searches for a unique match in the look-up table. If enough information is entered to locate a unique match, the software displays the match and prompts the interviewer to confirm that this is the correct entry. The interviewer can either accept the match or override the selection and continue to input the rest of the information. The second way is to enter the first few characters, then press "Enter" to bring up a list of all the entries in the look-up table that match the characters typed so far. The list appears in a pop-up window, and the interviewer can scroll through the list to select an entry. This feature is particularly useful when the information being entered is lengthy or difficult to spell, or if the interviewer expects a short list of matches. If no entry in the look-up table matches the typed characters, the software automatically displays a box in which the interviewer can type the rest of the name. A summary of these five browse-and-select windows is presented in Table 1, and examples are given in Figures 6 and 7.

The data for the look-up tables come from several sources. The municipality look-up table is based on a local listing from the Ministry of Transportation, Ontario. The street look-up table is an extraction from the Street Network File (SNF), which is managed by Statistics Canada. The SNF is similar to the TIGER file maintained by the U.S. Department of Commerce. The school listing is a result of a file on public and secondary schools from the Ontario Ministry of Education

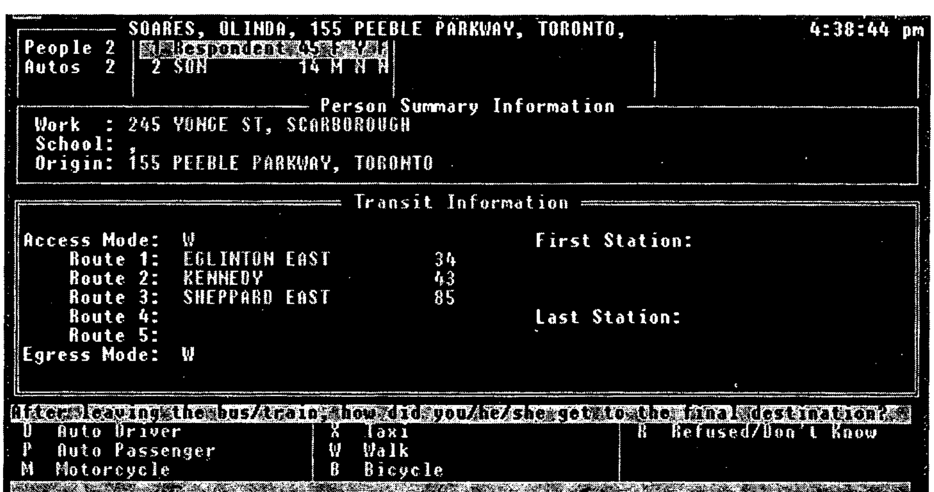

FIGURE 5 Transit data screen.

TABLE 1 Summary of Look-Up Tables

Type	Table Contents
Street	A listing of streets in the survey area sorted by name, type, direction and area municipality. Where there were two or more streets with the exact same description within a municipality, they were differentiated by hamlets. Selection of a street also determined the municipality name.
Municipality	In addition to the area municipality names, entries also included the names of local postal districts and hamlets.
School	All public, private, separate schools and post-secondary institutions in the survey area were included in this table. In addition, individual campuses were listed for all post-secondary institutions. The list was sorted by school name and area municipality.
Transit Route	Transit routes were searched by their formal name or alias name or route number. The list contained both the route and the operator names and included inter-city transit properties, as well as some privately operated shuttle services. Interconnections between transit routes were recorded in a separate database.
Transit Station	Commuter rail and subway stations were presented in separate listings.

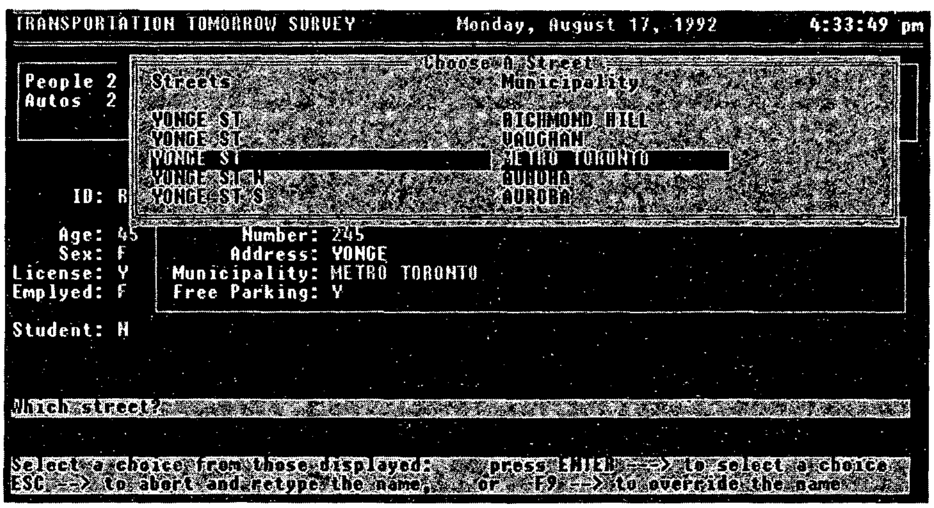

FIGURE 6 Street look-up table.

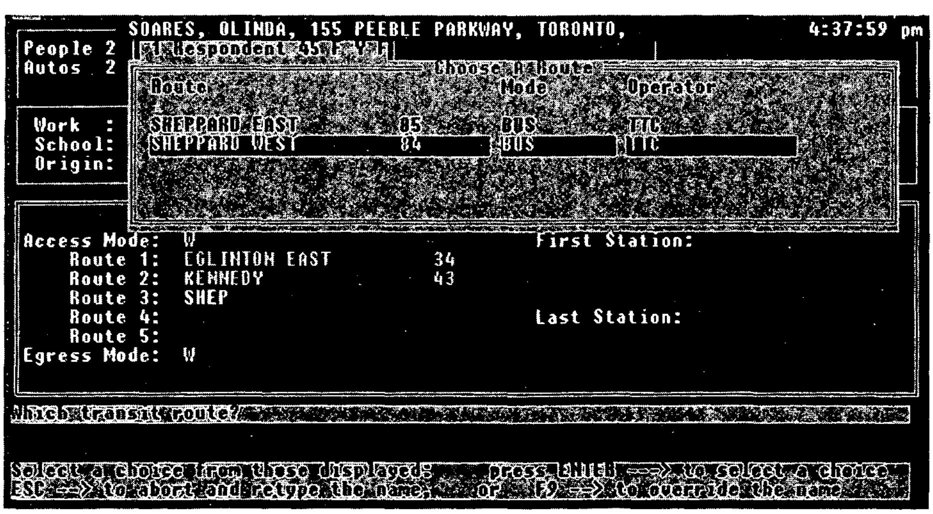

FIGURE 7 Transit route look-up table.

and a file on post-secondary schools collected from the individual regional municipality planning offices. Transit routes and station names are extracted from a transit modeling data base that is managed by the Data Management Group on behalf of the same agencies that sponsored the survey.

COMPARISON WITH 1986 SURVEY

To identify the costs and benefits of DDE, the 1991 survey (using DDE) and the 1986 survey (using paper and pencil) are compared extensively. The two surveys are nearly identical in survey area and questionnaire design.

Interview Statistics

A summary of the interview statistics is given in Table 2. When the DDE system was designed, it was not anticipated that the rate at which interviews could be completed would be increased over that achieved using the pencil-and-paper technique. The focus was on increasing the quality of the data and reducing the postprocessing requirements. In fact, the average number of completed interviews per paid hour of interview started at 3.0 and increased to about 4.0 at the end of the survey. Overall, the average rate was 3.76 interviews per hour. The better interviewers consistently recorded 4.5 to 5.0 completions per hour. These rates are higher than the rates achieved in 1986, when the average rate was 3.5 interviews per hour.

Monitoring procedures ensured that the high 1991 rates were not achieved at the expense of data quality. For quality control, some of the interviewers' computers were connected to monitoring screens accessible only to the survey supervisor. The supervisor was able to choose which interviewer to monitor by means of a master control panel. When using the panel along with a telephone monitoring system, the supervisor could visually monitor data being entered while listening to the interview. This direct monitoring system was especially useful during the interviewer training process. Interviewers with the highest completion rates also tended to have the fewest problems requiring call backs, and their trip rates per person were consistently uniform at or slightly above the overall average. The number of contacts required to complete a household interview was not tracked by itself. Instead, the combined number of call attempts (including unsuccessful call attempts) and post-interview edits was recorded. In total, 40 percent of the completed interviews required only one call/edit attempt, and another 40 percent were completed by the third call/edit attempt.

Several factors contributed to the exceptionally low refusal rate, which was 11 percent in 1991 (compared with 27 percent in 1986). Keeping this rate low was a function of the advance letter sent to inform the household about the survey, intensive interview training, smooth and quick responses of the DDE system, and the practice of keeping the interview short.

Despite the complexity of the travel survey and the amount of information collected, according to the telephone billing the average household interview took approximately 7 min to complete. This includes recording of basic household data (such as address, number of persons and vehicles), person information (such as age, gender, place of work, and place of school) and all trips made by household members 11 and older on a given weekday. Keeping this time short was again a function of the advance letter (which confirmed the legitimacy of the survey), the interviewer script built into the DDE system, intensive interviewer training, and the efficiency of the software in generating look-up tables and cross referencing the data.

The average household size observed in the 1991 survey was 2.8 persons, the same as in 1986. The trip rate increased slightly from 2.4 trips per person in 1986 to 2.6 in 1991. In terms of survey content, the 1986 and 1991 surveys collected almost identical information. The major difference was that a person's usual place of work or school was collected, if applicable, in 1991 but not in 1986.

Postprocessing Statistics

One of the goals of the survey was to ensure that the geocoding process kept pace with the conduct of the interviews. In that way, any problems could be corrected immediately and call backs could be made while the interview was still fresh in the

TABLE 2 Interview Statistics

	1986 TTS	1991 TTS
Number of interview stations (computers and telephones for 1991)	75	33
Number of staff recruited (interviewers and supervisors)	225	70
Average completion rate (interviews per paid hour)	3.50	3.76
Number of households in study area	1,466,000	1,716,600
Population in study area	4,063,000	4,730,000
Sample used (attempted to contact)	102,606	34,167
Eligible contacts (excluding wrong numbers, numbers out of service, etc.)	83,764 (82%)	27,813 (81%)
Refusal rate (%)	27	11
Completed interviews	61,453	24,507
Invalid or unusable interviews	255	146
Households in final data base	61,453	24,507
Overall completion rate (%)	60	72

minds of the respondents. This objective was achieved by a team of six geocoders and one supervisor, a ratio of 1:5 to the number of interviewers on an average interview night. In 1986 the geocoder-to-interviewer ratio was about 1:2. With few exceptions all location information was geocoded within 3 working days of the actual interview, and the entire coding operation was fully completed within 1 month of the end of the survey. In 1986 it took an additional 6 months after the completion of the survey to fully code all of the information.

The success of the coding operation may be credited partly to the high quality control throughout the conduct of the survey and partly to the efficiency of the geocoding software. Geocoding was straightforward when survey data were complete and accurate as a result of the on-line look-up tables in the DDE software. In fact, 48 percent of all location data were batch-geocoded by matching the surveyed data with a geographic information system (GIS) data base. The GIS data base included listings of postal codes, street names with address ranges, intersections, and major monuments (activity centers) such as hospitals, educational institutions, shopping malls, and large employment generators. The monument file was created by including listings from various government ministries (such as Health and Education) and monument files from other survey projects. Only 7 percent of the interviews required call backs to clarify ambiguous location descriptions. One of the most time-consuming tasks in the 1986 survey was the coding of transit route information, which had to be done entirely manually. In 1991 only 5 percent of this information needed to be manually coded, a significant savings in postprocessing effort.

Cost Comparisons

A comparison of the 1991 survey variable costs with the variable costs of the 1986 survey is given in Table 3. An inflation factor of 27 percent, the rise in the consumer price index, has been used to adjust the 1986 cost to 1991 values. The unit cost of interviewing was 30 percent higher than in 1986. The higher cost was primarily due to the significantly higher wages paid to recruit and retain good interviewers and the need for individual computers and the electronic monitoring equipment. The higher interviewing costs were more than offset by the reduction of 70 percent in unit coding cost, a direct result of better quality control in the conduct of the survey by using DDE and improvements in the geocoding procedures. The other variable costs remained basically unchanged. The small saving in the printing and mailing of the advance letter is due to the higher overall completion rate in 1991 (72 versus 60 percent), resulting in fewer letters having to be mailed relative to the number of interviews completed. The total variable costs for the 1991 survey, at $12.69/completed interview, are estimated to represent a 19 percent saving relative to the cost of the 1986 survey.

Comparing the fixed costs associated with the two surveys is more difficult because of the difference in management structures between the surveys. In 1986 a significant amount of staff time and ancillary support was contributed by the agencies for whom the survey was performed. This included staff assistance in the management of the survey, training interviewers, software development, and use of office space and computers. These costs were never specifically accounted for. In 1991 the costs for all of these tasks and resources were accounted for by the survey managers (the Data Management Group).

Another difference between the surveys was the size. Economies of scale were achieved in 1986 since three times more interviews were completed than in 1991. As a result, the documented 1991 development costs were significantly higher, at about $7.00/interview as opposed to $0.80/interview in 1986. Despite the high development cost, the overall cost of the 1991 survey was about $27.00/completed household interview. The fact that DDE is a relatively new technology to travel surveys meant that its development cost was expected to be high initially, but it should be reduced substantially in future applications.

TABLE 3 Variable Cost Comparison Between 1986 and 1991

	Adjusted 1986 Cost per Household	1991 Cost per Household
Interviewing		
Interviewers & Supervisors	$ 4.73	$ 6.29
Equipment & Supplies	$ 1.81	$ 2.20
	$ 6.54	$ 8.49
Coding		
Data Entry	$ 1.44	n/a
Geocoding/Transit coding	$ 5.41	$ 2.03
	$ 6.85	$ 2.03
Other Variable Costs		
Advance Letter	$ 1.23	$ 0.85
Other Direct Costs	$ 1.09	$ 1.33
	$ 2.33	$ 2.18
Total Variable Cost	**$15.72**	**$12.69**
Per Person	$ 5.67	$ 4.29
Per Trip	$ 2.62	$ 1.98

CONCLUSIONS AND RECOMMENDATIONS

The DDE system as implemented for the 1991 TTS was found to be a success. The system was effective in

- Maintaining high standards of quality control,
- Enabling a higher-than-expected rate of successful interview completion,
- Reducing dramatically the effort required for postprocessing the survey data, and
- Achieving all of this in a cost-effective manner.

Its use in other surveys, either telephone or home interview, can be recommended with the following provisos:

- Other parts of the process, including sampling selection, performance monitoring, and coding, should be automated as part of a totally integrated process.

- Adequate lead time must be available for development, testing, and interview training on the use of the software.
- Support staff, with the appropriate computer skills, must be available for trouble shooting at all times during the conduct of telephone surveys.

ACKNOWLEDGMENTS

This paper is based on the conduct of the 1991 TTS sponsored by the regional municipalities of Durham, Halton, Hamilton-Wentworth, Peel, and York; the municipality of metropolitan Toronto; the Ministry of Transportation, Ontario; GO Transit; and the Toronto Transit Commission. The design and conduct of the TTS was managed by the Data Management Group, Joint Program in Transportation, University of Toronto. The DDE software was written by BA Consulting Group Ltd.

REFERENCES

1. Steuart, G. N., B. Mori, and P. Noehammer. Direct Data Entry Using Microcomputer—A Travel Survey Pilot Project. In *Transportation Research Record 1134*, TRB, National Research Council, Washington, D.C., 1988, pp. 40–48.
2. Rukholm, T. *Design and Conduct of the Survey, 1986 Survey Series*. Toronto Area Transportation Planning Data Collection Steering Committee, Ministry of Transportation of Ontario, Downsview, Ontario, Canada, Dec. 1987.
3. Tranplan Associates. *Trip Diary Survey Analysis*. Ministry of Transportation of Ontario, Downsview, Ontario, Canada, Jan. 1990.
4. Dalton, P. D., and J. C. N. Ng. *Design and Conduct of the Survey, 1991 Survey Series*. Toronto Area Transportation Planning Data Collection Steering Committee, Joint Program in Transportation, University of Toronto, Ontario, Canada, Oct. 1992.

The views and conclusions expressed in this paper are those of the authors and not necessarily those of the participating agencies including the Data Management Group and BA Consulting Group Ltd.

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

Survey Approach for Study of Urban Commuter Choice Dynamics

HANI S. MAHMASSANI, THOMAS JOSEPH, AND RONG-CHANG JOU

A methodology to capture the day-to-day dynamics of user behavior in a commuting context is described. A two-staged survey was designed to obtain detailed information on the commuting habits in the north Dallas area for an extended period of time. In the first stage of the survey a one-page, two-sided questionnaire was sent to 13,000 households in the selected area. Information from the first stage provided a reliable characterization of the population of interest and prevailing commuting patterns. The second stage involved respondents of the first stage who were willing to provide more information on their commuting habits. This stage was considerably more detailed, consisting primarily of an activity diary limited to the commuting trips from home and returning to home. It included information on trip chaining, departure time, and path choice, at a level of detail previously unavailable. The information was obtained for a period ranging from 1 to 2 weeks. The two-staged format proved to be a cost-effective and practical method of obtaining the kind of information needed to study the dynamics of commuting behavior.

Several emerging policy concerns in transportation planning place significant new requirements on the understanding of travel behavior, on the ability to predict it, and consequently on the information available to characterize it and model its various aspects. In particular, efforts toward congestion mitigation, driven in part by concern over air quality attainment and energy efficiency considerations, have generated considerable activity in the area of travel demand management in its various forms. Strategies such as telecommuting, peak spreading actions through flexible hours, and increased vehicle occupancies require information on aspects of travel behavior that go far beyond that available through conventional travel surveys. Similarly, the opportunities for better system operation through information technologies require deeper understanding of the behavior of trip-makers. Central to the successful development and implementation of these strategies and to the attainment of related policy objectives is the consideration of the users' responses over time, which requires the characterization of current choices as well as an understanding of the underlying behavioral decision processes. Of particular concern to congestion-related strategies are work commuting trips, which continue to account for the notorious a.m. and p.m. peak periods in most major urban areas and as such are the primary target for the kind of strategies mentioned earlier.

There is only a limited observational basis on the dynamics of trip-makers' decisions, as these affect their responses to new policies. Commuter decisions are central to peak formation and evolution, and therefore to efforts aimed at reducing or spreading the peak. Virtually no information is available on the daily fluctuations of user decisions and of the resulting flows. Four principal travel choice dimensions are key determinants of those phenomena: trip chaining, trip timing, path choice, and modal choice (including carpooling), with modal choice probably taking place over a longer time frame than the first three. In addition, interactions among these dimensions need to be considered, preferably in context of the pattern of activities in which commuters are engaged.

It has become clear that traditional approaches to planning data acquisition, primarily in the form of cross-sectional home or phone interview surveys documenting a single day of travel, provide only limited information to address the kind of phenomena central to emerging policy concerns. Single-day cross-sectional travel surveys are inadequate as a basis for policy analysis and for studying the essential travel behavior processes. Longitudinal data are required for this purpose, at a level of detail normally unavailable in travel surveys, especially with regard to trip timing and path selection decisions. Pas (*1*) and Pas and Koppelman (*2*) have illustrated the importance of daily variation of travel choices and advocated the use of multiday surveys on both substantive and statistical grounds. Kitamura (*3*) has provided a thorough review of activity-based approaches to travel behavior analysis, indicating their importance to transportation policy analysis and highlighting the need for the kind of data for these studies.

Kitamura and Van der Horn (*4*) and Kitamura and Bovy (*5*) along with other coworkers have highlighted the kind of behavioral and policy issues that can be addressed with longitudinal data and proposed methods to deal with specific methodological issues that arise in connection with such data, with particular reference to the Dutch Panel study. A systematic discussion of the uses of different types of travel survey data for various functional needs in transportation decision making has recently been presented by Taylor et al. (*6*), though the special requirements associated with detailed study of short-term dynamics of commuter behavior are not addressed. Several earlier review papers on travel survey methodologies are available in the literature, such as those by Brog et al. (*7*) and Stopher (*8*).

The dynamics of commuter decisions in congested corridors have been the subject of laboratory-like experiments conducted by Mahmassani and coworkers (*9–11*). These interactive experiments had real commuters supply departure time and route choices in a simulated traffic system. Such experiments provide a useful observational basis for insights into the underlying behavioral processes, but they may not necessarily correspond to the commuters' actual settings. Such experiments must also typically introduce simplifications in

Department of Civil Engineering, University of Texas, Austin, Tex. 78712.

order to retain sufficient experimental control and avoid overly complex response tasks. For example, trip chaining was not considered in these experiments, even though it is significant in the commuting context as noted by Hanson (*12*) and Oster (*13*). For this reason, observation of commuters in their actual daily commutes is the necessary next step beyond laboratory experiments, as discussed by Mahmassani and Herman (*11*).

The data required to study commuter decision processes in actual commuting are detailed, requiring specific information on times of departure and arrival and intermediate stops and detailed link-by-link descriptions of the paths followed. Such data are not usually available in conventional travel surveys. For this reason, a survey approach was developed to obtain information for the study of commuter behavior dynamics. It consists primarily of an activity diary limited to the commuting trips from home and returning to home. The survey provides a unique level of detail, especially with regard to path choice, for which all links used in each reported commute are listed. In addition, trip chaining information in the form of the location, duration, and purpose of stops made along the commute are reported. The times of all events (departure, stop arrival, stop departure, arrival) are also reported. This information was obtained from commuters for a period ranging from 1 to 2 weeks, with several waves conducted in the Dallas, Texas, area.

The objectives of this paper are to (a) describe the survey approach developed and document its implementation; (b) share the methodological and substantive insights from two such surveys, a small one in Austin, Texas, and a more extensive one in Dallas; and (c) illustrate the kind of information on commuter behavior that can be obtained in surveys of this type. Although the approach was intended primarily for the study of commuter behavior dynamics, it provides a useful foundation for a survey procedure with more general applicability.

SURVEY DESIGN AND IMPLEMENTATION

The survey addressed in this paper was conducted in conjunction with a study on user responses to traffic disruptions and control strategies (*14*). The main objectives were to obtain information on the factors that affect trip-related decision making and capture the day-to-day variability of departure time and route choice of commuters. In addition to the substantive insights, the data were intended to develop user decision making models for a comprehensive day-to-day dynamic framework for the analysis and evaluation of traffic control strategies (*14*).

A study of trip-related decision making and factors associated with it is difficult in a large network due to the many travel options available to the traveler. Since much work-oriented travel tends to take place in a corridor context, the study is focused on a commuting corridor, which is simpler and more convenient for survey and analysis purposes than a general network. Selection of the study area considered the following characteristics:

- The majority of the work trips should terminate in a zone within the study area, and
- The area should contain distinct major facilities that anchor the principal commuting routes (e.g., freeways or major arterials) that are parallel to each other and terminate in a zone within the study area.

In the study area in Dallas, located north of the central business district (CBD), west of the North Central Expressway (Highway 75), and east of the Dallas Tollway (Figure 1), most work-related trips terminated in the CBD. Several parallel facilities passed through or terminated in the CBD: Dallas Tollway, Preston Road, Hillcrest Road, Coit Road, Greenville Avenue, Skillman Road, Abrams Road, and the North Central Expressway.

In addition, major reconstruction was also scheduled along the North Central Expressway around the time of the survey. It was hoped that the survey would therefore also provide data on the adjustment behavior of commuters during a long-term disruption.

The type of information required for the study posed a unique challenge in the design of the survey methodology. The necessary information on traffic, socioeconomic, and workplace characteristics of the commuters could be obtained from routine one-time survey questionnaires, but such questionnaires cannot provide information on the dynamic aspects of commuter behavior. The latter would require data on trip and traffic characteristics over several days. The detailed nature of the data makes it difficult to obtain the information reliably through retrospective-questioning surveys. Taylor et al. (*6*) and Duncan et al. (*15*) have addressed the inadequacies of these techniques, which include the participants' inability to recall details of especially the short-distance trips. They advocated the use of diaries, which could capture 15 to 40 percent more trips than conventional recall procedures. These methods are time-consuming and costly. Participants must be dedicated in order to participate reliably in the study. A one-time bulk mailing of the required number of questionnaires to obtain the desired response rate would have been very inefficient and expensive. A two-stage survey format through the mail was designed for the purpose. The survey was designed along the same lines as one conducted in the Austin area in 1989 (*16*). The first stage consisted of a one-page questionnaire wherein socioeconomic and commuting characteristics of the participants were sought. It also served as a screening device for participants of the second stage. The screening was based on the respondents' willingness to provide information as well as their willingness to participate in the second stage. The second stage consisted of a trip diary sent to selected candidates from the first stage who were willing to provide more information. Both stages were conducted through the mail. The following sections describe the design and implementation of the two stages.

First Stage

The first-stage survey questions were designed to achieve the following objectives:

- Acquire data on items that are relatively constant over extended periods (e.g., commuter characteristics),

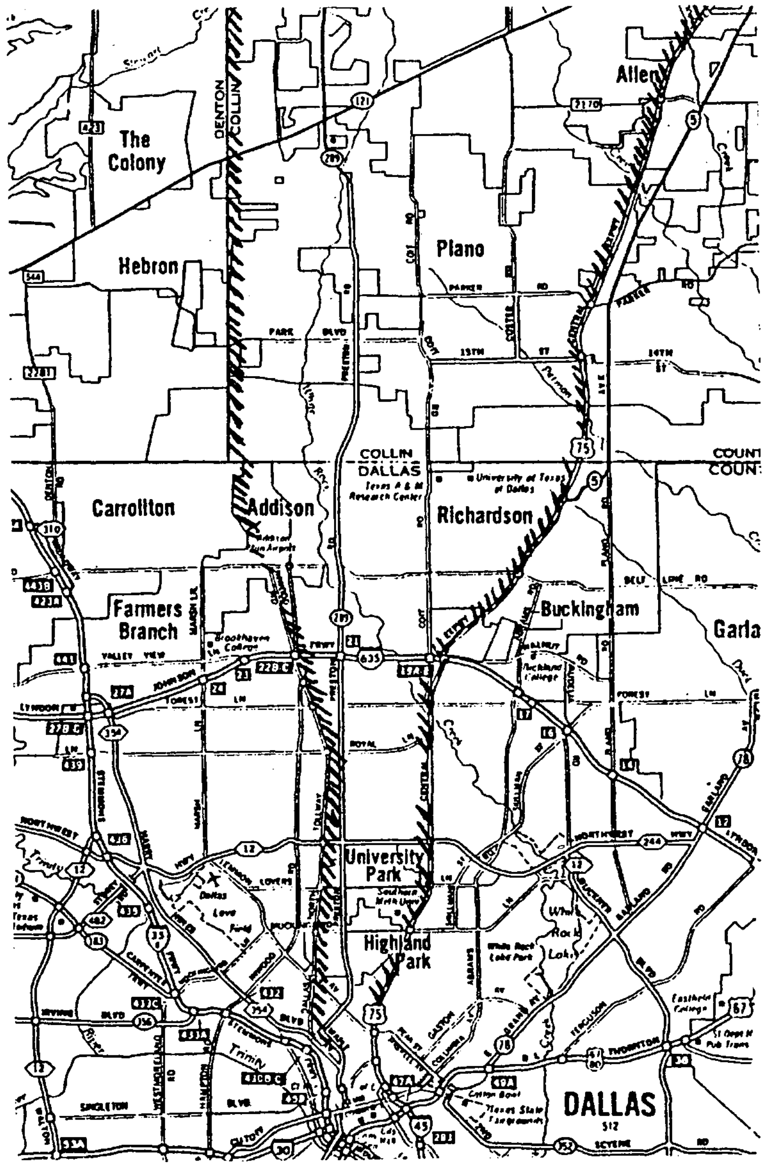

FIGURE 1 Survey area.

- Obtain information on commuter attitudes and other potentially important factors that contribute to the decision-making process, and
- Provide a mechanism to screen for prospective candidates for the second stage.

Questions in this survey can be split into three categories. The first category addressed the first survey objective and included questions on the workplace address, mode of travel to work, type of work, and usual commuting time. Responses to these questions were expected to remain constant during the survey period and used to characterize the commuter trip-making situation. This information was also used in screening and sampling candidates with the desired characteristics for the second stage.

The second category of questions addressed commuter attitudes and other decision factors, including the following:

- Decision state—whether the behavior of the commuter appears routinized, limited problem solving, or extensive problem solving with respect to trip-related decision making (e.g., the question asking the commuter if he or she normally adjusts the departure time or route specifically with traffic conditions in mind;
- Decision mediators—the factors that affect the decision (e.g., the commuter's attitude toward the various factors affecting route choice, such as number of signals or safety);
- Information acquisition process—whether the commuter actively or passively acquires information for trip-related decision making (e.g., whether the commuter owns a cellular phone or normally obtains information on traffic conditions before or during the trip); and
- Evoked set of alternatives—the possible alternatives that the commuter considers during trip-related decision making (e.g., the frequency of use of the various routes).

A question with significant implications on commuter behavior asked for the commuters' preferred arrival time (PAT) at work, which was found to be an important determinant of the dynamics of commuter behavior in previous experiments (*17*). This question was subject to different interpretations by the commuters. For example, a commuter may have had an initial PAT that was unattainable in the current situation because of congestion or parking problems. The commuter may have reconciled to another attainable PAT, which would have been reported in response to the question. Two versions of the question were designed. About half the households were asked to provide the PAT with no conditions set (Case 1), and the other half were asked to provide it under the assumption of no congestion and parking difficulties (Case 2). Analysis of the distributions of the PAT obtained from both versions indicated statistically significant differences between the two cases, with a higher fraction of commuters indicating a PAT of 0 in Unconstrained Case 1. However, in both cases, about 50 percent of respondents indicated a preference for arrival at the workplace within 10 min before the official work start time (*18*). A related question asked how important it was for the commuter to be on time for work. The response to this question would reflect the combined effects of the actual policy at the work place, the perception of the policy by the commuter, and the personal characteristics (attitude) of the commuter toward arriving late for work. Interestingly, more than a third of the respondents indicated unlimited lateness tolerance and more than half reported no lateness tolerance (*18*).

The third category of questions addressed the socioeconomic characteristics of the commuters that may be related to commuter behavior. These included job title, owning or renting a home, number of children, and the like. Although these questions were primarily used to generate sample demographics, it was also expected that some of these variables would bear on certain aspects of commuter behavior.

A final question asked if the participant was willing to provide more detailed information on his or her commuting status. Three responses were possible: yes, no, and possibly. The "possibly" option was included to retain potentially agreeable commuters who were not yet willing to commit without obtaining information. This option was included after the previous experience from the survey in Austin suggested that there was a considerable pool of candidates who were not willing to participate in the second stage simply because they were not sure of what was expected of them. A comparison of the fraction of positive responses from this survey with those from the Austin survey clearly established the advantages of including this option.

Implementation of First Stage

Several steps were taken to promote the professional appearance, comprehensibility, and user-friendliness of the questionnaire, which consisted of a two-sided, single letter-size sheet. The questions were all numbered and grouped into major sections in a logical pattern and laid out to maximize readability. A code number was used instead of the respondents' names to preserve the confidentiality of responses. Before the questionnaires were mailed a pilot survey was conducted at the University of Texas, Austin, to determine if changes needed to be made for clarity and proper interpretation of the questions. No significant problems were identified.

The survey area comprised nine ZIP code zones and encompassed the major part of the North Central Expressway and its alternative routes. A sample size of 13,000 households was selected from the area. The number of households randomly selected from each zone was proportional to the population within the zone. The 13,000 households were further split into two groups, each receiving one of the two versions of the questionnaire.

Each of the selected households was sent a packet in the mail consisting of a cover letter, two questionnaires (same version), and a self-addressed business return envelope. The cover letter described the seriousness of traffic congestion and discussed ongoing research to develop new and innovative solutions to curb the problem at the University of Texas. The purpose was to generate interest in the study and hopefully increase the response rate. The University of Texas letterhead and envelopes were used to convey a sense of sincerity and importance and to distinguish it from junk mail.

Second Stage

The second stage provided the central information desired of the survey, namely, information on individual daily decision making and the variation of departure and route choice decisions. Commuters were asked to provide information on their trip characteristics for several days. The implementation costs per participant and the effort on their part were considerably more than in the first stage. The screening process in the first stage was intended to identify a pool of interested and willing candidates.

Two types of diaries, a long and short version, were designed to record the day-to-day behavior of the participating commuters over a 2-week period. The duration of the trip diary (10 working days) was considered sufficient for examining short-term dynamic behavior but not so long as to harm the respondents' goodwill.

The second stage differed from the first stage in that the amount of interpretation and recollection was reduced while the level of detail was greatly increased. Data from this stage were expected to be more accurate than those from the first stage. On the long version of the diary, the commuter was asked to record, for each trip to work, the departure time, target arrival time, actual arrival time, official work start time, link-by-link details of route selected, and the time, duration, and type of intermediate stops made. The target arrival time was meant to indicate the commuters' predicted time of arrival at work. Information on the number, type, and nature of the stops during commutes should provide insights into the importance of trip chaining and extent of pretrip planning. Insights into the commuter prediction and decision-making process could be obtained from the target arrival time, extent of early/late arrivals, and the associated changes made to trip schedules subsequently. It was realized that different commuters might interpret the question on the target arrival time differently. For example, commuters may consider the target time as the time at which they were required to arrive at their workplace rather than a consequence of some sort of travel time prediction process.

If reconstruction activity was observed during a particular trip, the commuter was asked to note down the street along which this occurred. Two questions were directed toward the acquisition of information on traffic conditions before and during the trip. From the responses to these, valuable information on the extent of pretrip planning, states of commuter decision making (e.g., routinized, extensive problem solving, etc.), and the potential for information-based strategies (e.g., ATIS/ATMS) can be extracted. Commuters were also asked to indicate if they had observed any accidents or traffic jams during their trip.

The level of detail required in route description was significantly different in the long and short versions of the diary. In the long version, a link-by-link description of the route including minor deviations was required, but only the name of the major facility used along the commuting route was requested in the short version. Similarly, commuters were asked for the details of every intermediate stop in the long version (the arrival and departure times and the purpose of the stop); only the number of stops was required in the short version.

Questions pertaining to the trip from work were similar to those for the trip to work. At the end of the survey, commuters were asked to respond to six final questions on the last page of the diary. The first three questions were related to parking and included the type of parking, cost of parking, and time to travel from the parking lot to the workplace. The last three questions were related to information acquisition and measured the propensity to acquire and use information if provided and the potential of various information sources. A detailed description of the questions is found elsewhere (*14*).

Implementation of Second Stage

The primary screening criterion for the second-stage survey was based on the response to the question, in the first-stage survey, on whether the respondents were willing to provide additional information on their commuting habits. In the first stage, 2,658 (10.22 percent) total responses were obtained. Of the 2,521 useful responses, 1,249 indicated "yes," 804 indicated "possibly," and 468 indicated "no" in response to the participation question. The advantage of including the "possibly" option is clearly indicated by the greater fraction of probable participants than in the Austin survey (about 80 percent in this survey, compared with about 55 percent in the Austin survey) (*16*). Only commuters responding with a "yes" or "possibly" in the first stage were considered for the second stage. From this set of respondents, those with unsuitable characteristics (e.g., retired, very short travel time, frequent out-of-state travel, work at home, no fixed work location, walk or bike to work) were deleted. The diaries were mailed during the first and second weeks of April 1990.

As explained earlier, the maximum duration for the participation of a given commuter was limited to 2 weeks. To obtain information on commuter patterns in the area over a longer period during the initiation of the freeway reconstruction activity, two overlapping subwaves involving different participants were administered. The first subwave extended from June 11 to June 22, 1990. The second subwave extended from June 18 to June 29, 1990. A second wave was conducted about a year later in an attempt to capture any long-term effects of the reconstruction activity on commuter patterns (this paper does not address the results of the second wave). Participants in the second wave included a combination of new participants and participants who had taken part in the previous wave. To enhance the response rate on the diaries, telephone calls were made to a considerable number of prospective participants at strategic times to retain their participation.

The sample for the first wave was divided into four groups on the basis of the type of work of the participants. About half the participants were sampled randomly from these groups for each subwave. Adjustments were made to the final sample to ensure that two eligible candidates from the same household were always grouped together to participate in the same subwave of the survey.

Four hundred candidates who responded with a "yes" to the participation question were sent the long version of the diary, on the hypothesis that commuters who respond "yes" rather than "possibly" are likely to be more committed to participating and would therefore be more likely to fill out the longer versions of the diary. The remaining section of the sample (which included a few remaining "yes" responses and all the "possibly" replies) were sent the shorter version. Again, two eligible members of the same household always received the same version of the diary regardless of their response to the participation question. If for example, one of the two participants of the household was selected to receive the longer version, the other would automatically be included in the list and sent a long version even if he or she responded with "possibly" to the participation question. Figure 2 and Table 1 show the sampling strategy and sample details for the second stage.

Each of the selected households received a package consisting of a cover letter; one or two diaries (same version), depending on the number of participants; and a reply envelope. The cover letter thanked the respondents for their participation in the first stage and described in brief what was expected during the second stage. A phone number was also provided for questions that participants may have before or during the survey period.

The diary booklets were 8.5 in. long and 3.5 in. wide. They were constructed so as to easily fit into the return envelopes (which was a size smaller than the mailing envelopes) provided without folding. The booklet was designed to be convenient for the commuter to handle while in the car. Each day had separate predated pages for the morning and evening commutes. Also included in the booklets were detailed instructions and a sample of a completed day's entries. The cover of the diary was made of thicker material to protect the contents from excessive wear and tear. Codes rather than return addresses were also used in this stage. Figures 3 and 4 illustrate sample pages from the short and long diaries, respectively. The decision to use stamped instead of business reply envelopes for this stage was based on considerations of the mailing costs per package and the expected response rates.

Of the 2,053 willing candidates, those with commuting characteristics that were unsuitable for the survey were eliminated. This resulted in 1,973 eligible participants. During the first

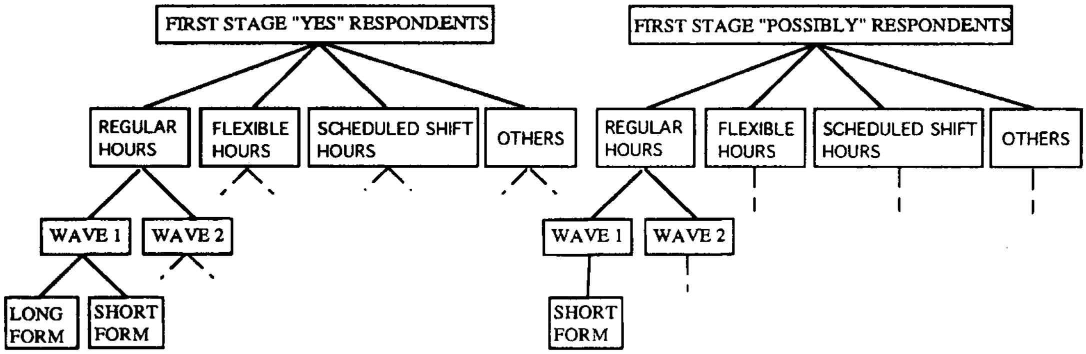

FIGURE 2 Sample page from diary (short version).

wave of the second stage, 783 long and 742 short versions of the diary were sent; 221 of the long and 231 of the short diaries were returned (Table 1). Not all of the returned diaries were useful for the study. Figure 5 displays the useful responses to Wave 1 on a daily basis. A mild trend toward a decreasing response rate is observed. The more significant trend is, however, the reduced response rates on all the Fridays of the survey duration. This is possibly because of the lower number of commutes made on Fridays or early weekends taken by certain participants.

SURVEY RESULTS

Several types of analysis were performed on the data obtained from the questionnaires and the diaries. First, an exploratory analysis of behavior with respect to the departure time and route choice and the stops was conducted. Frequency models (Poisson regression models) were also developed for departure and route switching and number of stops made. Finally, discrete choice models were developed to explain the day-to-day variations in the departure time and route choice. These models were of the multinomial probit type with a large number of alternatives and a general error structure. Comparisons were also made to the results from the earlier survey in Austin and those obtained from the laboratory-type experiments conducted at the University of Texas. A full report of all the analyses is found elsewhere (*18*). Table 2 presents a summary of characteristics from the first-stage survey. Some important observations made on the basis of the second-stage survey results are discussed in the following.

Only 75.1 percent of all morning and 64.1 percent of all evening commutes contain no stops at all, indicating that trip chaining is an essential feature of urban commuting. As expected, commuters stop more often during evening commutes, because of possibly less severe time constraints after work, as well as the availability of more stopping opportunities (more stores open, etc.). Although 32.2 percent of commuters did not make any stops on the way to work, only 18.6 percent never stopped on the way home. At the other extreme, only 6.8 percent of the commuters made stops on every morning trip and 5.8 percent of them made stops on every evening trip. The results indicate a wide spread of commuter trip-linking habits and daily variability in the commuting population.

Results of the departure time switching analysis indicate that commuters engage in a substantial amount of departure time switching for both morning and evening commutes. Departure time switching for evening commutes is more frequent than that for morning trips. One of the useful contributions

TABLE 1 Diaries Sent in Second Stage

		Short version		Long version	
Response to participation question	Type of work hour	subwave 1	subwave 2	subwave 1	subwave 2
Yes	Regular	65 (24)	59 (28)	262 (79)	294 (91)
Yes	Scheduled shift	5 (2)	2 (1)	11 (1)	4
Yes	Flexible	16 (3)	16 (4)	75 (20)	78 (19)
Yes	Other	3	3	9 (3)	7 (1)
Possibly	Regular	206 (59)	207 (76)	20 (3)	7 (2)
Possibly	Scheduled shift	13 (2)	8 (2)	1	1
Possibly	Flexible	58 (11)	63 (16)	5 (2)	2
Possibly	Other	9 (2)	4 (1)	2	2

Note: Number in parentheses is number of diaries returned.

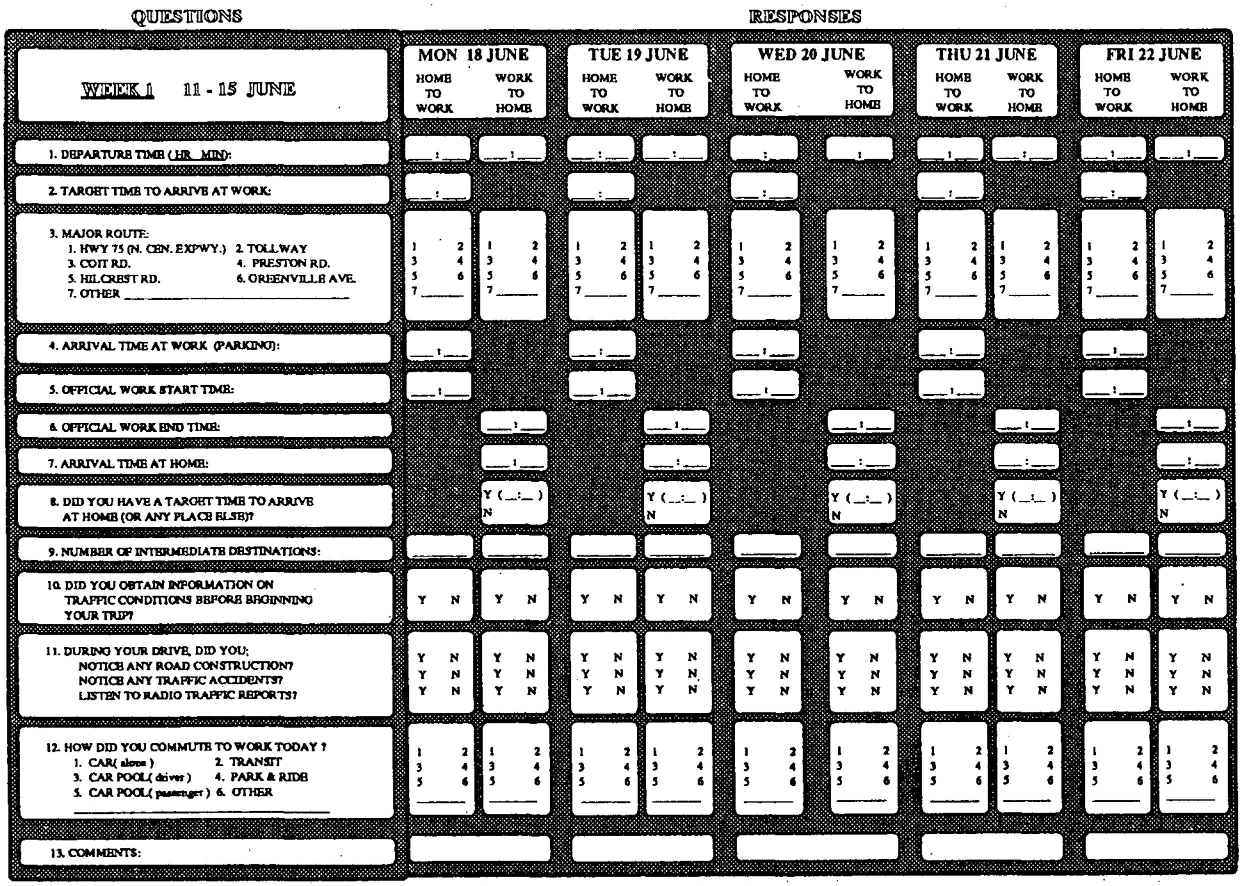

FIGURE 3 Sample page from diary (long version).

of this analysis is that it captures actual decisions of commuters in an uncontrolled environment, yielding a characterization of the natural variability of these decisions in a real system.

An analysis of the repetition and variability of the commuters' route choices during the 2-week survey period indicates that route switching is not as frequent as departure time switching for a.m. and p.m. commutes. Like departure time switching, route switching is more frequent during p.m. commutes than during a.m. commutes. The lower frequency of route switching relative to departure time switching is consistent with the results of stated preference experiments under simulated traffic conditions (*10*). These results suggest the potential of real-time information to influence the temporal distribution of trips to a greater extent than the spatial distribution of trips over the network routes.

Very little switching relative to the mode route occurs if only no-stop routes are considered, as 56.9 percent of the users never switch routes under these circumstances in the morning and 60.8 percent never switch routes in the evening. Clearly, the need to link one or more activities along the commute influences path selection and accounts for much of the variation in the selected routes.

CONCLUDING COMMENTS

This paper has presented a survey approach intended to capture the dynamic aspects of traveler behavior, specifically in connection with the home-to-work and work-to-home commuting trips. The survey information has yielded information on the extent of trip chaining associated with the commute and its variability from day to day. It has also documented the extent of daily fluctuations in the departure times for the commuting trip chains. The results suggest that the picture obtained from conventional single-day surveys of household trip making is incomplete and of limited use in connection with travel demand management and congestion mitigation strategies. The journey to work, considered among the more stable elements of urban travel demand, is itself variable from day to day and the magnitude of this variability is not insignificant, especially in connection with the aforementioned types of strategies. Similarly, the symmetry usually assumed between a.m. and p.m. trips is limited, with the p.m. commute subject to more variability than its a.m. counterpart.

Another unique feature of the survey is the level of detail of the information obtained, especially with regard to the selected paths through the network. Such information has been previously unavailable yet is of utmost relevance to current studies of electronic route guidance systems.

It is remarkable that commuters have generally been able to provide the information requested at the desired level of detail. The authors' analyses have uncovered only a relatively small number of inconsistencies in the responses, and follow-up contacts with the participants have confirmed some of the answers obtained and the participants' general comfort with the survey instruments. Considerable effort was invested on our part to ensure clear and user-friendly instruments.

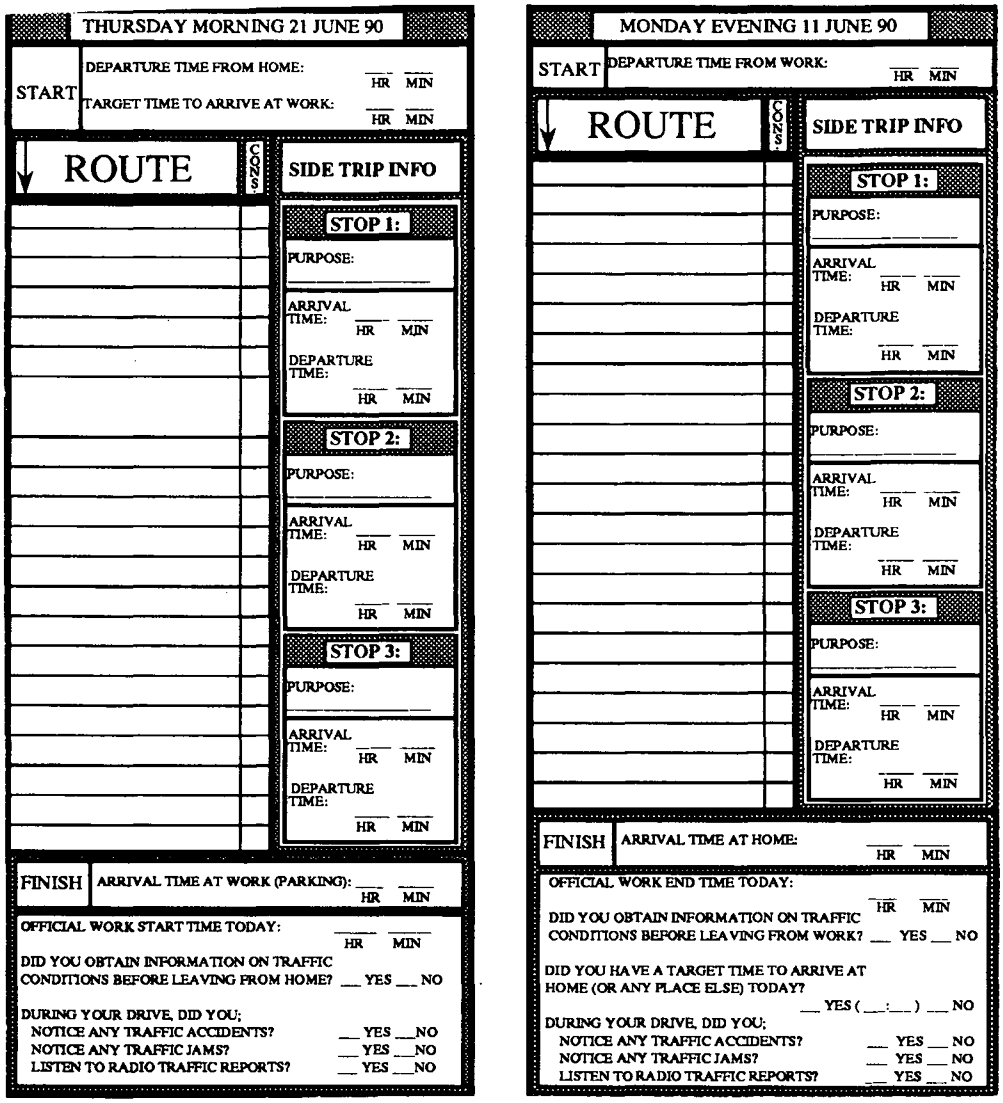

FIGURE 4 Sampling plan for second-stage survey.

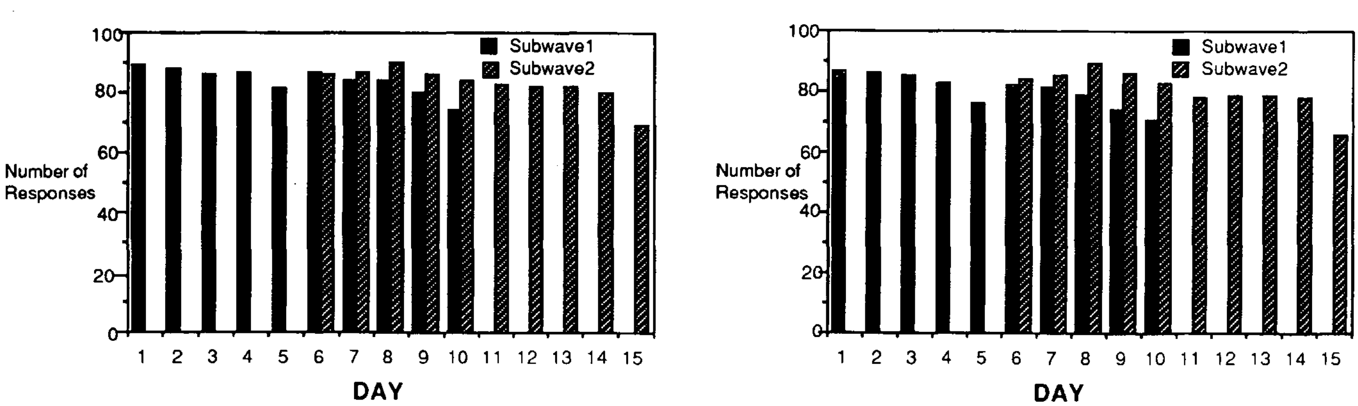

FIGURE 5 Daily responses for Wave 1 of Stage 2, a.m. (*left*) and p.m. (*right*).

TABLE 2 Summary Statistics for First-Stage Survey Results

Mode of Travel for Commuter (2518*)		
	Car(alone)	93.6%
	Car Pool	2.4%
	Transit	1.0%
	Park and Ride	1.3%
	Other	1.7%
Type of Work Hour (2518)		
	Regular Work Hours	70.5%
	Scheduled Shift Work	2.9%
	Flexible Work Hours	20.0%
	Other	6.6%
Preferred Arrival Time at Work Place		
	Case 1: No Conditions Specified (1178)	16 minutes
	Case 2: In the Absence of Congestion or Parking Problem (1192)	15 minutes
Tolerance to Late Arrival at Work Place (2492)		
	Unlimited	38.2%
	Given Time	7.3%
	None	54.5%
Average Daily Travel Time		
	From Home to Work(2485)	25 minutes
	From Work to Home(2346)	27 minutes
Commuter Adjusting Departure Time		
	From Home to Work(2489)	52.9%
	From Work to Home(2461)	31.4%
Commuter Modifying Route		
	From Home to Work(2487)	47.1%
	From Work to Home(2467)	46.1%
Arrival after Intended Time (2482)		
	More than Five Times	8.3%
	Between 1 and 5 Times	42.3%
	None	49.4%
Commuter Listening to Radio Traffic Report (2494)		70.6%
Commuter Having Cellular Car-Phone (2503)		10.5%
Age (2504)		
	Under 18	0.6%
	18-29	14.9%
	30-44	46.5%
	45-60	31.3%
	Over 60	6.7%
Gender (2505)		
	Male	63.3%
	Female	36.7%
Commuter Willing to Help Further (2514)		
	Yes	49.6%
	No	18.6%
	Possible	31.8%

* Total sample size is 2521. Value in parentheses is the number of responses for each question.

In retrospect, the short version of the second-stage survey was not as successful as anticipated. The response rate was not any higher for it than for the longer full diary. It would have been preferable to go only with the latter, which was done for the second survey wave.

Although the survey was intended for commuter trips, the insights gained suggest that the approach could be used to obtain a more complete record of trips and activities. The two-stage strategy was helpful in improving the cost-effectiveness of the second-stage survey by better targeting of households likely to yield usable responses. In addition, the first-stage survey yielded very useful information in its own right, in terms of providing a reliable characterization of the population of interest and prevailing commuting patterns. The nature of the questions in the first questionnaire and the elapsed time between the first and second stages provide interesting opportunities to contrast the diaries of actual behavior with previously reported responses. Such questions, along with the fundamental processes underlying the dynamics of traveler decisions, are the subject of ongoing and future work in connection with the rich observational basis obtained in this survey.

ACKNOWLEDGMENTS

The Dallas survey described in this paper is based on a survey initially conducted in Austin. Gregory S. Hatcher and Christopher G. Caplice were instrumental in the conduct and analysis of the Austin survey. Richard Rothery and Robert Herman contributed to the design of the survey instruments for the Austin survey as well. Chee-Chung Tong played a key role in both the Austin and Dallas surveys. Several individuals helped in the logistical aspects of the Dallas survey reported in this paper—in particular, Eileen Egan and Laura Brod played a major role in the coding of the diary information for computer analysis. The participation of Ray Derr, of the Texas Department of Transportation (TxDOT), in various stages of the survey is greatly appreciated. Funding for the work on which this paper is based was provided by TxDOT.

REFERENCES

1. Pas, E. I. Sample Design, Parameter Estimation Precision and Data Collection Costs for Least Squares Regression Trip Gen-

eration Models. Presented at the 64th Annual Meeting of the Transportation Research Board, Washington, D.C., 1985.
2. Pas, E. I., and F. Koppelman. An Examination of the Day-to-Day Variability in Individuals' Urban Travel Behavior. *Transportation*, Vol. 14, No. 1, 1986, pp. 13–20.
3. Kitamura, R. An Evaluation of Activity-Based Travel Analysis. *Transportation*, Vol. 15, 1987, pp. 9–34.
4. Kitamura, R., and T. Van der Horn. Regularity and Irreversibility of Weekly Travel Behavior. *Transportation*, Vol. 14, 1987, pp. 227–251.
5. Kitamura, R., and P. H. L. Bovy. Analysis of Attrition Biases and Trip Reporting Errors for Panel Data. *Transportation Research*, Vol. 21A, No. 4/5, 1987, pp. 287–302.
6. Taylor, M. A. P., W. Young, M. R. Wigan, and K. W. Ogden. Designing a Large Scale Travel Demand Survey: New Challenges and New Opportunities. *Transportation Research*, Vol. 26A, No. 3, 1992, pp. 247–262.
7. Brog, W., A. H. Meyburg, P. R. Stopher, and M. J. Wermuth. Collection of Household Travel and Activity Data: Development of a Survey Instrument. In *New Survey Methods in Transport* (E. S. Ampt, A. J. Richardson, and W. Brog, eds.), VNU Science Press, The Netherlands, 1985, pp. 151–172.
8. Stopher, P. R. The State of the Art in Cross-Sectional Surveys in Transportation. In *New Survey Methods in Transportation* (E. S. Ampt, A. J. Richardson, and W. Brog, eds.), VNU Science Press, The Netherlands, 1985, pp. 55–76.
9. Mahmassani, H., G. L. Chang, and R. Herman. Individual Decisions and Collective Effects in a Simulated Traffic System. *Transportation Science*, Vol. 20, No. 4, 1986, pp. 258–271.
10. Mahmassani, H., and D. Stephan. Experimental Investigation of Route and Departure Time Dynamics of Urban Commuters. In *Transportation Research Record 1203*, TRB, National Research Council, Washington, D.C., 1988, pp. 69–84.
11. Mahmassani, H., and R. Herman. Interactive Experiments for the Study of Tripmaker Behavior Dynamics in Congested Commuting Systems. In *Developments in Dynamic and Activity-Based Approaches to Travel Analysis* (P. Jones, ed.), Avebury, Aldershot, 1990, pp. 272–298.
12. Hanson, S. The Importance of the Multi-Purpose Journey to Work in Urban Travel Behaviour. *Transportation*, Vol. 9, 1980, pp. 229–248.
13. Oster, C. Second Role of the Work Trip—Visiting Nonwork Destinations. In *Transportation Research Record 728*, TRB, National Research Council, Washington, D.C., 1979, pp. 79–82.
14. Joseph T., H. Mahmassani, and R. C. Jou. *Day-to-Day Dynamics Framework for the Analysis of User Responses to Traffic System Disruptions and Control Actions*. Technical Report 1216-2F. Center for Transportation Research, University of Texas, Austin, 1992.
15. Duncan, G. F. The Role of Panel Studies in Research on Economic Behavior. *Transportation Research*, Vol. 21A, No. 4/5, 1987, pp. 249–263.
16. Caplice, C. *Analysis of Urban Commuting Behavior: Switching Propensity, Use of Information, and Preferred Arrival Time*. Master's thesis. University of Texas, Austin, 1990.
17. Mahmassani, H., and G. L. Chang. Dynamic Aspects of Departure-Time Choice Behavior in a Commuting System: Theoretical Framework and Experimental Analysis. In *Transportation Research Record 1037*, TRB, National Research Council, Washington, D.C., 1985, pp. 88–101.
18. Jou, R. C., H. Mahmassani, and T. Joseph. *Daily Variability of Commuter Decisions: Dallas Survey Results*. Technical Report 1216-1. Center for Transportation Research, University of Texas, Austin, 1992.

The views and findings expressed and described in this paper are the authors' and do not necessarily reflect those of TxDOT.

Publication of this paper sponsored by Committee on Travel Behavior and Values.

Underreporting of Trips in Telephone Interview Travel Surveys

MAZEN I. HASSOUNAH, LOY-SAI CHEAH, AND GERALD N. STEUART

Results of a research project on underreporting of trips in telephone interview household travel surveys arising from memory lapses and the use of proxies are presented. On the basis of a survey of approximately 61,000 households in the greater Toronto area, the effects of these two factors on reported automobile mobility characteristics were analyzed with respect to trip characteristics and to socioeconomic characteristics of households and individual trip makers. The analysis showed trip underreporting to be the rule for short discretionary trips and trips made during off-peak periods. Using these insights, correction procedures were developed to minimize the effects of trip underreporting.

In any travel survey, several types of bias or errors can be introduced or are inherent in the survey procedures. In general, there are two types of survey procedural biases (*1*):

1. Random sampling errors, which are introduced by the fact that a survey is a sample used to represent a population and influenced by factors such as sample size and method of sampling; and
2. Systematic errors, which are introduced by factors such as use of an incorrect sampling frame, insufficient control of sampling, excessive nonresponse, or consistent underreporting of trips.

The purpose of this paper is to draw attention to systematic biases in telephone interview travel surveys arising from consistent underreporting of trips by respondents.

Telephone interviews are a cost-effective method of collecting household travel information. They are easy to conduct and require fewer people to administer than other data collection techniques (*2*). However, research into travel survey methods has established that oral surveys in general and telephone interviews in particular produce relatively poor results in terms of trip reporting in comparison with written surveys (e.g., mail-back travel surveys) (*3–5*). Many factors have been proposed as sources of trip underreporting in telephone interview surveys. Prominent among these are use of proxies (i.e., informants report on trips made by third parties) and memory lapses (i.e., people forget to report trips). The research reported in this paper attempts to estimate the effects of these two factors on trip reporting and to develop correction procedures to minimize these effects.

EMPIRICAL DATA BASE

The empirical results of this paper are based on the 1986 Transportation Tomorrow Survey (TTS). The TTS was a telephone interview survey of approximately 4 percent of households in the greater Toronto area (GTA). The GTA is located on the northwest shore of Lake Ontario, Canada, and consists of metropolitan Toronto and the five regional municipalities of Durham, York, Peel, Halton, and Hamilton-Wentworth. The selection of telephone interviews as a method of collecting household travel information was motivated by the high telephone subscriber rate in the GTA and by the cost-effectiveness of such interviews. In 1986 there were approximately 1.5 million households in the GTA, and fewer than 2 percent of them were without a telephone.

The TTS was the first comprehensive areawide travel survey conducted in the GTA since 1964. The purpose of the survey was to collect household sociodemographic and travel behavior data that would be used in a variety of planning exercises. Representatives of households were asked to report on trips made by all members of the household during a prespecified weekday (usually the day before the interviewer call). Travel information such as origin location, destination location, purpose, mode, and start time of trips was collected. In addition, information was collected on household (location, dwelling type, size, and number of available private vehicles) and personal characteristics (age, gender, possession of driver's license, and employment status). When the TTS was concluded in December 1986, more than 61,400 households had been surveyed with information recorded for about 171,000 persons and 340,000 trips.

During the planning and implementation of the TTS, precautions were taken to avoid biases in the collected data. The random sample of households in the study area was selected from Bell Canada's residential billing files. The Bell files contain telephone numbers of households whose telephone numbers are listed in the telephone directory. Households without telephones or with unlisted numbers were found to be uniformly distributed throughout the study area with no obvious correlation with socioeconomic status. Sample households were assigned in a random fashion to individual interviewers to prevent any systematic variation in the quality of the interviews. Five attempts were made to contact each household. As a result, a high response rate of 73.7 percent was achieved. Data entry was accompanied by automated error checking (range checks and logic checks) with errors being referred back for correction.

The control totals used to expand the sample of households, persons, and trips to that of the total population in the GTA were based on a 268-zone system. The expansion was carried out on a household basis using 1986 census household information. To ensure spatial consistency of the expansion process, each zone was defined so as to contain at least 2,500

Joint Program in Transportation, University of Toronto, 42 St. George Street, Toronto, Ontario M5S 2E4 Canada.

household units reported in the census. For each household record in the TTS data base, an expansion factor was calculated as the ratio of the number of household units reported in the census to the number of surveyed household units in the aggregation zone where the household had been located. The same expansion factor was used for all trip and person records associated with the household. The average expansion factor in the TTS data base was 25. Comparisons of the expanded data with data from the census and other Statistics Canada surveys suggested that the TTS sample was generally representative of the GTA population in terms of household and population sizes and labor force participation rates.

Despite the precautions taken to ensure a high quality of response, systematic underreporting of trips was detected during the TTS data validation process. This underreporting was found to be severe for automobile trips, modest for transit trips, and almost negligible for walking and bicycle trips. Furthermore, underreporting of automobile trips was found to be a result of the use of proxies and memory lapses, as the following sections of this paper will show.

BIAS DUE TO USE OF INFORMANTS IN TTS

During the conduct of the TTS, proxy interviewing with any adult member of the household was adopted on the understanding that the respondent, while being interviewed, would ask other members of the household, if available, particulars on their trips. To examine the quality of response (i.e., trip rates reflecting trip recall), informant and noninformant trip rates were compared generally during the data validation process of the TTS. (An informant is an individual who reported his or her own trips as well as trips made by other members of the household, and a noninformant is an individual whose trips were reported by somebody else in the household.) This simple analysis revealed a significant difference ($2.703 - 1.854 = 0.849$ trips per person) in the overall trip rate of informants and noninformants. This difference in trip rate can be due to the informants' incomplete knowledge of trips made by other members of their households or to differences in the characteristics of the two groups (i.e., informants and noninformants). Informants having incomplete knowledge of trips made by noninformants leads to the underreporting of the noninformants' trips. If this underreporting is different for different kinds of trips or different groups of people, a bias is introduced to the data.

To investigate the effect of the use of informants on reporting of trips and the resulting bias, if any, a number of analyses were performed on the reported TTS trip rates of informants and noninformants. First, a descriptive analysis of TTS trip rates of informants and noninformants by various trip characteristics (purpose, length, and the time of day the trip was made) and socioeconomic characteristics of trip makers (gender, age, possession of driver's license, employment status, place of residence, household size, dwelling type, and the number of vehicles available for household members) was performed to identify probable factors that might contribute to the difference in the reported average number of trips of informants and noninformants. The factors that were identified as probable contributors to the difference included age and household size of trip makers in addition to all trip characteristics considered in the analysis. Consequent analyses of variance were then performed to determine whether these factors show statistical significance in contributing to the difference in the trip rate of informants and noninformants.

The results of these analyses indicated the difference in trip rate of informants and noninformants to be inconsequential for home-based-work/school (HBWS) trips (Table 1) (trips from home to work or school or vice versa) and significant for home-based-discretionary (HBD) and non-home-based (NHB) automobile short trips occurring outside the morning peak period of 6:00 to 9:00 a.m. (Tables 2 and 3). (An HBD trip is a trip from home to a destination other than work or school or vice versa, and an NHB trip is a trip that neither originated nor terminated at home.) Furthermore, the difference between informant and noninformant HBD trip rates was found to vary significantly across households of different sizes (Table 2). This data bias, however, was independent of the time of day that the trip was made (Table 3).

On the basis of these findings, procedures for correcting the effect of the informant were developed and applied differentially to the subsets of TTS trips that had been found to have data bias. The procedures were based on correction factors that incorporated the ratios of informant to noninformant trip rates. A summary of the estimated correction factors is given in the following table:

Trip Purpose	*No. Persons in Household*	*Factor*
HBD	2	1.404
	3	2.142
	4 or 5	2.780
	> 5	3.625
NHB	n/a	3.134

The factors were applied to TTS trip data in the same manner that TTS expansion factors were applied. The exception was that the correction factors were applied as multipliers to trip records that match the trip or household characteristics as defined in the previous paragraph and in the previous table.

TABLE 1 Analysis of Variance of HBWS Trips

	Probabilities Calculated F Values are Exceeded	
Stratification	HBWS	Others
Age	0.0001	0.0001
Household Size	0.0456	0.0882
Respondent Status (Informant/Non-informant)	0.4999	0.0001
Age by Respondent Status	0.0020	0.0295
Household Size by Respondent Status	0.8724	0.0112

TABLE 2 Analysis of Variance of HBD and NHB Trips

Trip Length (km)	Stratification	Probabilities Calculated F Values are Exceeded	
		HBD	NHB
< 5	Respondent Status	0.0008	0.0008
	Household Size by Respondent Status	0.0001	0.0463
	Age by Respondent Status	0.0059	0.1809
5-25	Respondent Status	0.0538	0.0448
	Household Size by Respondent Status	0.8431	0.4143
	Age by Respondent Status	0.6458	0.9169
25-50	Respondent Status	0.2697	0.8168
	Household Size by Respondent Status	0.7738	0.7137
	Age by Respondent Status	0.9894	0.9357
> 50	Respondent Status	0.0101	0.0460
	Household Size by Respondent Status	0.0335	0.3449
	Age by Respondent Status	0.2436	0.5529

For instance, an HBD automobile short trip (i.e., less than 5 km in straight line distance) made outside the morning peak period of 6:00 to 9:00 a.m. by a noninformant from a household of two persons was multiplied to 1.404 trips.

Application of the use-of-informant correction procedures to the TTS data resulted in an increase of approximately 34 percent in the number of expanded TTS daily automobile trips. Despite this increase, comparisons of corrected TTS travel data with selected cordon line counts in the GTA showed TTS automobile trips to still be underreported.

CORDON COUNT EVIDENCE OF TRIP UNDERREPORTING IN TTS

Even when a travel survey is conducted with meticulous care to avoid biases and trip underreporting, it is almost impossible to have complete agreement between survey trip data and cordon counts (*5*). One reason is that travel surveys are always subject to random sampling errors introduced by the fact that a survey is a sample used to represent a population and is influenced by sample size and method of sampling. Another reason is that travel surveys and cordon counts are usually carried out at different times and, unless measures are taken to account for this discrepancy, they are temporally incompatible. Unlike cordon counts, estimated cordon crossings derived from survey data do not include crossings made by nonresidents of the study area or by taxis and service vehicles. Cordon counts can also be overestimated because of multiple cordon crossing trips, and cordon line counts are subject to seasonal variation and may vary from day to day because of such factors as weather, road construction, and traffic accidents. Finally, it should be noted that a respondent trip log in a travel survey is not a factual log of trips but the respondent's recall of his or her travel activities, which may not be complete or accurate because of memory lapses.

Table 4 presents estimates of automobile person trip underreporting as calculated from cordon count and TTS automobile travel data at selected cordon lines in the GTA. Estimates of TTS automobile person trips crossing the cordon lines were obtained by running user equilibrium and all-or-nothing assignments of TTS automobile trips, corrected for the use-of-informant effect, in peak and off-peak periods, respectively. During any one of the considered time periods, variations in the calculated trip underreporting rate at the different cordon lines are relatively small, which suggests that underreporting of TTS trips occurs uniformly in space. It is evident from Table 4 that TTS automobile person trips were systematically underreported throughout the day. The extent of underreporting of trips, however, varies significantly from one period to another. It is modest in the morning peak period of 6:00 to 9:00 a.m., slightly worse in the evening peak period of 3:00 to 6:00 p.m., and worst in the midday off-peak period of 9:00 a.m. to 3:00 p.m.

The following table presents percentage shares of Metro's automobile trips by purpose in different periods during the day as estimated from TTS data:

Period	*HBWS*	*HBD*	*NHB*
6 to 9 a.m.	75	17	8
9 a.m. to 3 p.m.	24	50	26
3 to 6 p.m.	51	31	18

TABLE 3 Analysis of Variance of Trips by Starting Time

Stratification	Probabilities Calculated F Values are Exceeded			
	6:00 to 9:00	9:00 to 15:00	15:00 to 18:00	18:00 to Midnight
Respondent Status	0.0171	0.0001	0.0019	0.0001
Age by Respondent Status	0.0041	0.2744	0.3737	0.0206
HHLD Size by Res. Status[a]	0.0317	0.3196	0.0608	0.1836

[a] Household Size by Respondent Status.

TABLE 4 Cordon Count Evidence of Trip Underreporting

	Under-Reporting Rate[a]		
Cordon Line	6:00 to 9:00	9:00 to 15:00	15:00 to 18:00
Metro-York	-3.0%	-44.8%	-8.3%
Metro-Durham	-2.6%	-45.7%	-10.2%
Metro-Peel	-2.8%	-47.5%	-9.7%
York-Durham	-3.1%	-46.8%	-10.7%

[a] Rate=((TTS assigned volume/cordon count)-1)*100.

On the basis of this table and the findings reported in the previous paragraph, some conclusions can be deduced about the relation between underreporting of trips and trip purpose. First, the table indicates that HBWS trips dominate urban trip making during the morning peak period, whereas HBD and NHB trips dominate urban travel during the midday off-peak period. Given that the extent of underreporting of trips in general was found to be modest in the morning peak period and worst in the midday off-peak period, one can conclude that HBWS trips were much better reported in TTS than HBD and NHB trips were. Furthermore, the magnitude of underreporting of HBWS trips appears to remain constant throughout the day, whereas that of HBD and NHB appears to vary from one period to another. This is suggested by the relatively small rate of trip underreporting in the evening peak period in which the amount of HBWS trip making is almost equal to that of HBD and NHB.

The cordon count data used in this analysis were available only for the periods described earlier. The evening off-peak time period of 6:00 p.m. to 12:00 a.m. was, therefore, excluded from the analysis. In developing procedures to correct underreporting, underreporting of trips in this period was assumed to be analogous to that of the period 9:00 a.m. to 3:00 p.m. In addition, analysis of underreporting of HBD trips as distinct from NHB trips could not be facilitated by the data and, consequently, the two trip purposes were assumed to be underreported in a similar manner.

PROCEDURE FOR CORRECTION OF TRIP UNDERREPORTING EFFECT IN TTS

The basic idea of the procedure was to increase the level of TTS automobile trips, disaggregated by trip purpose and corrected for the use-of-informant effect, up to that of cordon line counts by means of time-dependent correction factors. The factors were developed by comparing TTS peak and off-peak automobile person trips in each purpose category reported as crossing the Metro Toronto boundary with those obtained from cordon line counts. This was facilitated by assuming trips of all purposes to be equally underreported during the morning peak period of 6:00 to 9:00 a.m. and the HBWS trip underreporting rate to remain constant throughout the day. Both assumptions were justified on the basis of the results of the previous section.

The complete procedure for correction of the trip underreporting effect in TTS consists of the following steps:

1. Estimate the morning peak-period trip underreporting correction factor from TTS data and cordon line counts,
2. Apply this factor to HBWS trips during other periods of the day, and
3. Estimate HBD and NHB trip underreporting correction factors during other periods of the day on the basis of the results of Steps 1 and 2 and on percentage shares of automobile trips as reported in the previous table.

A summary of the estimated correction factors is as follows:

Period	*HBWS*	*HBD*	*NHB*
6 to 9 a.m.	1.03	1.03	1.03
9 a.m. to 3 p.m.	1.03	1.85	1.85
3 to 6 p.m.	1.03	1.09	1.09
6 p.m. to 12 a.m.	1.03	1.85	1.85

It is noteworthy that cordon line counts may not be free of underreporting as implied throughout the outlined correction procedure. Because of the usual siting of cordon lines along natural boundaries, only trips that are long enough to cross these boundaries may be counted, whereas short and localized trips may go unrecorded. As a consequence, HBD and NHB travel, which consists mainly of short and localized trips, may remain underreported even after corrections.

CONCLUSIONS

This paper has demonstrated the extent of trip underreporting in telephone interview travel surveys due to memory lapses and use of proxies. The effects of these two factors on trip reporting were analyzed with respect to a number of trip and socioeconomic characteristics. The analysis showed trip underreporting to be the rule for short discretionary trips. The analysis also showed that trips made during off-peak periods are more likely to be underreported than trips made during peak periods. On the basis of these findings, correction procedures were developed to account for trip underreporting.

Several steps can be taken to reduce the potential for trip underreporting in telephone interview travel surveys. For instance, the scope of each household interview can be broadened to include all members of the household. Such a measure, however, should be balanced against costs and respondent response rates. The use of direct data entry software can also help in reducing trip underreporting in telephone surveys. Such software can have features that check trip connectivity and consistency as the data are being collected. This allows the interviewer to query the respondent when any gap appears in the trip log of any household member.

REFERENCES

1. Roberts, M., and H. Gunn. Response Errors and Data Correction. In *New Survey Methods in Transport* (E. S. Ampt, A. J. Richardson, and W. Brog, eds.), VNU Science Press, The Netherlands, 1985, pp. 28–31.
2. Meyer, M., and E. J. Miller. *Urban Transportation Planning: A Decision-Oriented Approach.* McGraw-Hill Book Company, New York, 1984.
3. Brog, W., and H. Meyburg. Consideration of Nonresponse Effects in Large-Scale Mobility Surveys. In *Transportation Research Record 807*, TRB, National Research Council, Washington, D.C., 1981, pp. 39–46.
4. Meyburg, H., and W. Brog. Validity Problems in Empirical Analyses of Non-Home-Activity Patterns. In *Transportation Research Record 807*, TRB, National Research Council, Washington, D.C., 1981, pp. 46–50.
5. Barnard, P. Evidence of Trip Underreporting in Australian Transportation Home Interview Surveys and Its Implications for Data Utilizations. In *New Survey Methods in Transport* (E. S. Ampt, A. J. Richardson, and W. Brog, eds.), VNU Science Press, The Netherlands, 1985, pp. 311–331.

Publication of this paper sponsored by Committee on Traveler Behavior and Values.