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

Cordon Line	Under-Reporting Rate[a]		
	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.