UNDER-REPORTING OF TRIPS IN TELEPHONE INTERVIEW TRAVEL SURVEYS

Data Management Group
Joint Program in Transportation
University of Toronto

January 1993

UNDER-REPORTING OF TRIPS IN TELEPHONE INTERVIEW TRAVEL SURVEYS

By

Mazen I. Hassounah, Loy-Sai Cheah and Gerald N. Steuart

Joint Program in Transportation, University of Toronto

42 St. George Street, Toronto, Ontario, M5S 2E4, Canada

Tel. (416) 978-7282 Fax (416) 978-3941

ABSTRACT

This paper presents results of a research project on under-reporting of trips in telephone interview household travel surveys arising from use of proxies (i.e. use of informants to report on trips made by third parties) and memory lapses (i.e. people forgetting to report trips). Based on a survey of approximately 61,000 households in the Greater Toronto Area, Ontario, Canada, the effects of these two factors on reported auto mobility characteristics (average number of auto trips per person, total number of auto trips) were analyzed with respect to trip characteristics (purpose, length, start time) and with respect to socioeconomic characteristics of households and individual trip makers. The analysis showed trip under-reporting to be the rule for short discretionary trips and trips made during off peak periods. Utilizing these insights, correction procedures were developed to minimize trip under-reporting effects.

INTRODUCTION

In any travel survey, several types of bias or errors can be introduced, or are inherent to the survey procedures. In general, there are two types of survey procedural biases (1):

(i) random sampling errors 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

(ii) systematic errors introduced by factors such as use of an incorrect sampling frame, insufficient control of sampling, excessive non-response, or consistent under-reporting of trips.

The purpose of this paper is to draw attention to systematic biases in telephone interview travel surveys arising from consistent under-reporting 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 relative to 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, 4, 5). Many factors have been proposed as sources of trip under-reporting in telephone interview surveys. Prominent among these are use of proxies (i.e. use of informants to report on trips made by third parties) and memory lapses (i.e. people forgetting 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% 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. Fewer than 2% of households were with no telephone.

The TTS was the first comprehensive area-wide travel survey conducted in the GTA since 1964. The purpose of the survey was to collect household socio-demographic and travel behaviour 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 pre-specified weekday, i.e. usually the day previous to the interviewer call. Travel information such as origin location, destination location, purpose, mode and start time of trips was collected. In addition, information about households such as location, dwelling type, size and number of available private vehicles, and personal information like age, sex, possession of driver's license and employment status was collected. When the TTS was concluded in December, 1986, over 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 results, a high response rate of 73.7% 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 a minimum of 2500 household units reported in the census. For each household record in the TTS database, 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 database 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 labour force participation rates.

In spite of the precautions taken to ensure high quality of response, systematic under-reporting of trips was detected during the TTS data validation process. This under-reporting was found to be severe in the case of auto trips, modest in the case of transit trips, and almost negligible in the case

of walk and bicycle trips. Further, under-reporting of auto 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), general comparisons of informant (i.e. an individual who reported his/her own trips as well as trips made by other members of the household) and non-informant (i.e. an individual whose trips were reported by somebody else in the household) trip rates were performed during the data validation process of the TTS. This simple analysis revealed a significant difference (2.703 - 1.854 = 0.849 trips/person) in the overall trip rate of informants and non-informants. This difference in trip rate can be due to:

(i) informants having incomplete knowledge of trips made by other members of their households, and/or

(ii) differences in the characteristics of the two groups (i.e. informants and non-informants).

Informants having incomplete knowledge of trips made by non-informants leads to the latter's trips being under reported. If this under-reporting is different for different kinds of trips and/or different groups of people, a bias is introduced to the data.

In order 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 non-informants. First, a descriptive analysis of TTS trip rates of informants and non-informants by various trip characteristics (purpose, length, and the time of day the trip was made) and socioeconomic characteristics of trip makers (sex, age, possession of driver's licence, 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 non-informants. 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 non-informants. The results of these analyses indicated the difference in trip rate

of informants and non-informants to be inconsequential in the case of home-based-work/school (HBWS) trips (Table 1) (a HBWS trip is a trip from home to work/school or vice versa), and significant in the case of home-based-discretionary (HBD) and non-home-based (NHB) auto short trips occurring outside the morning peak period of 6:00 to 9:00 a.m. (Tables 2,3) (a HBD trip is a trip from home to destinations other than work/school or vice versa, and a NHB trip is a trip that neither originated nor terminated at home). Further, the difference between informant and non-informant 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 the trip was made (Table 3).

Based on these findings, procedures for correction of the use-of-informant effect 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 non-informant trip rates. A summary of the estimated correction factors is given in Table 4. 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 and/or household characteristics as defined in the previous paragraph and in Table 4. For instance, a HBD auto short (i.e. < 5 km in straight line distance) trip made outside the morning peak period of 6:00 to 9:00 a.m. by a non-informant from a household of size 2 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% in the number of expanded TTS daily auto trips. In spite of this increase, comparisons of corrected TTS travel data with selected cordon line counts in the GTA showed TTS auto trips to still be under-reported.

CORDON COUNT EVIDENCE OF TRIP UNDER-REPORTING IN TTS

Even when a travel survey is conducted with meticulous care to avoid biases and trip under-reporting, 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 influenced by sample size and method of sampling. Another reason is that travel surveys and cordon counts are usually carried out in different time periods and, unless measures are taken to account for this discrepancy, they are temporally incompatible. Further, unlike cordon counts, estimated cordon crossings derived from survey data do not include crossings made by non-residents of the study area, and by taxis and service vehicles. Furthermore, cordon counts can also be over estimated due to multiple cordon crossing trips. In addition, cordon line

counts are subject to seasonal variation and may vary from day to day as a result 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 (i.e. the recall) may not be complete or accurate due to memory lapses.

Table 5 presents estimates of auto person trip under-reporting as calculated from cordon count and TTS auto travel data at selected cordon lines in the GTA. Estimates of TTS auto person trips crossing the cordon lines were obtained by running user equilibrium and all-or-nothing assignments of TTS auto 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 under-reporting rate at the different cordon lines are relatively small which suggests that under-reporting of TTS trips occurs uniformly in space. Further, it is evident from Table 5 that TTS auto person trips were systematically under reported throughout the day. The extent of under-reporting of trips, however, varies significantly from one time period to another. It is modest in the morning peak period of 6:00 to 9:00, slightly worse in the evening peak period of 15:00 to 18:00, and worst in the mid-day off peak period of 9:00 to 15:00.

Table 6 shows percentage shares of Metro's auto trips by purpose in different time periods during the day as estimated from TTS data. Based on this table and the findings reported in the previous paragraph, some conclusions can be deduced about the relation between under-reporting of trips and trip purpose. First, the table shows HBWS trips to dominate urban trip making during the morning peak period, while HBD and NHB trips to dominate urban travel during the mid-day off peak period. Given that the extent of under-reporting of trips in general was found to be modest in the morning peak period and worst in the mid-day off peak period, one can conclude that HBWS trips had been significantly better reported in TTS than HBD and NHB trips. Further, the magnitude of under-reporting of HBWS trips seems to remain constant throughout the day, while that of HBD and NHB appears to vary from one time period to another. This is suggested by the relatively small rate of trip under-reporting 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 the above analysis were available only for the time periods described above. The evening off peak time period of 18:00 to 24:00 was, therefore, excluded from the analysis. In developing procedures for the correction of under-reporting, under-reporting of trips in this period was assumed to be analogous to that of the time from 9:00 to 15:00. In addition, analysis of under-reporting 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 under-reported in a similar manner.

A PROCEDURE FOR CORRECTION OF THE TRIP UNDER-REPORTING EFFECT IN TTS

The basic idea of the procedure was to increase the level of TTS auto 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 auto 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 under-reported during the morning peak period of 6:00 to 9:00, and the HBWS trip under-reporting rate to remain constant throughout the day. Both assumptions were justified based on the results of the previous section.

The complete procedure for correction of the trip under-reporting effect in TTS consists of the following steps:

(i) estimation of the morning peak period trip under-reporting correction factor from TTS data and cordon line counts,

(ii) application of this factor to HBWS trips during other periods of the day, and

(iii) estimation of HBD and NHB trip under-reporting correction factors during other time periods of the day based on the results of steps (i) and (ii) and on percentage shares of auto trips as reported in Table 6.

A summary of the estimated correction factors is given in Table 7.

It is noteworthy that cordon line counts may not be free of under-reporting as implied throughout the outlined correction procedure. Due to the usual siting of cordon lines along natural boundaries, only trips that are long enough to cross these boundaries may be counted, while short and localized trips may go unrecorded. As a consequence, HBD and NHB travel, which consists mainly of short and localized trips, may remain under-reported even after corrections.

CONCLUSIONS

This paper has demonstrated the extent of trip under-reporting 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 under-reporting to be the rule for short, discretionary trips. The analysis also showed that trips made during off peak periods are more likely to be under-reported than trips made during peak periods. Based on these findings, correction procedures were developed to account for trip under-reporting.

Several steps can be taken to reduce the potential for trip under-reporting in telephone interview travel surveys. For instance, the scope of each household interview can be broaden 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 (DDE) software can also help in reducing trip under-reporting in telephone surveys. The DDE 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 any household member's trip log.

REFERENCES

1. M. Roberts and H. Gunn. Response Errors and Data Correction. In New Survey Methods in Transport, E. S. Ampt, A. J. Richardson and W. Brog (Eds.), pp. 28-31. VNU Science Press, The Netherlands, 1985.

2. M. Meyer and E. J. Miller. Urban Transportation Planning: A Decision-Oriented Approach. McGraw-Hill Book Company, New York, 1984.

3. W. Brog 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. H. Meyburg 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. P. Barnard. Evidence of Trip Under-Reporting 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.), pp. 311-331. VNU Science Press, The Netherlands, 1985.

TABLE TITLES

TABLE 1 ANALYSIS OF VARIANCE OF TTS TRIPS (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 TTS TRIPS (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

TABLE 3 ANALYSIS OF VARIANCE OF TTS TRIPS (BY START TIME OF TRIPS)

	Probabilities Calculated F Values are Exceeded			
Stratification	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 CORRECTION FACTORS (FOR BIAS DUE TO USE OF INFORMANTS)

Trip Purpose	Household Size	Factor
HBD	2 persons	1.404
	3 persons	2.142
	4-5 persons	2.780
	> 5 persons	3.625
NHB	n/a	3.134

TABLE 5 CORDON COUNT EVIDENCE OF TRIP UNDER-REPORTING

	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.

TABLE 6 PERCENTAGE SHARES OF AUTO TRIPS IN METRO TORONTO

Time Period	HBWS	HBD	NHB
6:00 to 9:00	75	17	8
9:00 to 15:00	24	50	26
15:00 to 18:00	51	31	18

TABLE 7 FACTORS FOR CORRECTION OF THE TRIP UNDER-REPORTING EFFECTS

Time Period	HBWS	HBD	NHB
6:00 to 9:00	1.03	1.03	1.03
9:00 to 15:00	1.03	1.85	1.85
15:00 to 18:00	1.03	1.09	1.09
18:00 to Midnight	1.03	1.85	1.85