PATTERNS OF PROXY USAGE IN THE 2001 NATIONAL HOUSEHOLD TRAVEL SURVEY

Jonaki Bose, and Lee Giesbrecht
Bureau of Transportation Statistics, 400 7th Street, SW, Room 4432, Washington DC 20590

Keywords:
Proxy Responses, Travel Data, NHTS, Data Quality

Introduction to the 2001 National Household Travel Survey (NHTS)

The 2001 National Household Travel Survey (NHTS) is the first comprehensive household survey of both daily and long distance travel, allowing for analysis of personal travel by Americans. The NHTS is intended to be used by researchers and analysts in metropolitan planning organizations, universities, and state and local governments, to understand how people make travel and residential decisions and predict their impact on public transportation from national and local perspectives.

The 2001 NHTS data include a nationally representative sample of 60,000 individuals in 26,000 households. The sample included all members of the household, including children under the age of five and individuals with disabilities. The NHTS used a random digit dial (RDD) telephone sampling design. 2001 NHTS data were collected from March 2001 to May 2002. Data were collected about all household members either directly from the respondent or through a proxy. A household where 50 percent of the adults completed the survey was considered a responding household and included in the data file. Individuals from sampled households were asked to complete a travel diary documenting their daily trips in order to aid in the recall process, and use the diary when responding to the interviewer. The overall response rate was 41 percent. (The household screener interview rate was 58 percent, and the useable household rate was 71 percent.)

The survey included items on the numbers of drivers and vehicles in the household, the characteristics of these vehicles, and the driver status of all people residing in the household, along with their views on transportation, use of public transit, walking, and biking activities in the last week. In addition, for each individual, data were collected on daily and long distance trips taken during pre-assigned time frames. For daily trips, data were collected on trip times, means of transportation, which household vehicle was used, if any, wait and access and egress (time from end of transit trip to final destination) times for transit trips, purpose, and presence of household and non-household members.

For long distance trips of 50 miles or more one way, information was collected on the number of trips made during a four week period, dates for the trips, whether the trip was recurring, purpose and destination of trip, type of lodging used at the destination, primary means of transportation, overnight stops, and access and egress information on airplane, bus, and train trips. In addition demographic information such as age, sex, medical condition status, and country of birth was collected on the people residing in the household.

Use of Proxies in the NHTS

Proxies are used in a number of federal surveys including the National Crime and Victimization Survey, the Survey of Income and Program Participation, and the National Health Interview Survey.

In the NHTS, an adult household member always served as the proxy for a child under age 14. Proxies were also requested for persons age 14 and 15 years. However, if an adult household member requested that the interviewer speak directly with these teenagers, the interview was conducted with the subject. Proxies were not initially requested for household members 16 years and older, but were allowed under limited conditions. Respondents with proxy data were strongly encouraged to complete their travel diaries. Proxy interviews were conducted for 22.5 percent of the respondents 18 and older, and 26 percent of respondents in households with more than one adult (table 1). Essentially,

in almost half of the households with more than one adult, a single adult reports for all other household adults (table 2). Despite the emphasis on a completed diary for proxy respondents, only 67 percent of proxy respondents had a completed diary compared to 64 percent of persons who responded for themselves (table 3). It is not clear who filled these diaries out—the sampled person or the person responding on their behalf. One of the reasons proxies are used in the NHTS is that data about a respondent's randomly assigned travel day have to be collected within six days, and this relatively short window for data collection makes it harder to interview all household members.

Prior Research

There is not a large body of work examining the effects of proxy responses in the context of travel surveys. There are however, a few surveys that are relevant to our efforts. In an experimental study of the effects of level of participation on proxy reports of vacation planning (Bickart, Blair, Sudman, and Menom 1991), proxies tended to underreport their partner's behavior and proxy reports about activities were the least accurate of the measures they examined. These findings are similar to results from a study on adult education participation in the adult education supplement of the CPS (Collins, Brick, Kim, and Stowe, 1997). Proxies tended to underreport the categories of: any activity, any activity excluding full time college, and any activity excluding college. In a proxy study of the Survey of Program Dynamics, they found that proxy respondents did not know work schedules for the other respondents in the household (Hess, Rothgeb, and Zukerberg 1997). Badoe and Steuart (2002) in a study of the effect of proxy respondents in the 1996 Transportation Tomorrow Survey of the Greater Toronto Area found that home-based discretionary and non home-based trips were underreported by proxy and that gender was a related factor. Kojetin and Miller (1993) found in a study of the Consumer Expenditure Surveys that parents were poor proxies for children's spending behavior. Since parents report for virtually all children in the NHTS, this finding may apply to travel behavior as well. Based on these findings, we were concerned that proxy responses could have led to an underreporting of travel behavior in the NHTS.

Research Questions and Analysis

We developed the following research questions:

- Are adults with proxy responses demographically different from self-reporting adults?
- Do key travel statistics differ for these two groups?
- What factors are related to obtaining a proxy report instead of a self report?

Correspondingly, the analysis is divided into three parts.

1. First, we compared self reports to proxy reports on the following basic demographic characteristics: sex, age, education, driver and worker status, country of birth, presence of a medical condition affecting travel, and whether the respondent was away from the home during the randomly assigned day that daily travel was collected about. In addition, key travel statistics such as mean daily trips, mean long distance trips, mean weekly bike trips, mean weekly walk trips and percentage having used public transit on the travel day and in the last two months were calculated for both proxy and self respondents. Though these differences were calculated for all households and households with more than one adult, this paper presents these tables only for households with more than one adult.

2. In the second part, we examine whether differences in key travel statistics persist even when controlling for demographic differences (using linear and logistic regression models).

3. In the final part of the analysis, we look at factors related to the likelihood of being a proxy respondent (using demographic characteristics in a logistic regression-based model).

It is very difficult, however, to evaluate the extent of under- or over-reporting without a true

experimental set up. In truth, we have no 'real' data on proxies. In order to conduct this analysis we had to consider certain data to be true. We postulated that the following demographic variables would not adversely suffer from a proxy effect: sex, age, education, driver and worker status, country of birth, presence of a medical condition affecting travel, and whether the respondent was away from the home during the randomly assigned daily travel. We are aware that this assertion may not be entirely accurate.

Even though we did not automatically ask an adult to answer for 16- and 17-year olds, a large number of adults did respond for their children. This study is restricted to adults over 18. Many of these results also pertain to households with more than one adult, because these are the households where there technically could have been a proxy respondent.

Results and Discussion

Adults with proxy responses do differ in terms of their demographic and travel characteristics. In terms of bivariate characteristics, higher percentages of proxy respondents are male, are between 18 and 24 years in age, have a high school diploma or less, are non-drivers, are foreign-born, have a medical condition affecting travel, and were away from home on their randomly assigned travel day compared to self respondents (table 3). There also are differences in key travel statistics. Self-reporting persons indicate taking more daily, long distance, bike and walk trips and a greater proportion report using public transportation (table 4). Some of these differences persisted even when looking at trip types (table 5). These differences in key travel statistics between proxy and self respondents continue to persist after controlling for the demographic differences mentioned before (tables 6 and 7).

Certain variables are more strongly related to the likelihood of an adult needing a proxy respondent in a multivariate context (table 8). Men are over twice as likely as women to have proxy responses. Non-drivers were also more likely (1.96 times) to have a proxy response. Persons away from the home on the assigned travel day, with a high school diploma or less or with a medical disability affecting travel were also less likely to be self-respondents. There are a few differences in key factors related to whether a man or woman would be a proxy respondent (table 9). The biggest difference lay in the fact that foreign-born women were 1.7 times more likely to have a proxy response than US-born women, while U.S-born men were slightly more likely to have a proxy response (1.3 times more likely) compared to foreign-born men.

Overall, we do not know if these are true differences or due to measurement error. For example, are proxy respondents more likely to take fewer trips since they are more likely to be non-drivers or have medical disabilities affecting travel, and thus more likely to be home bound? Or are they out and about traveling and thus likely to make more trips without the knowledge of the person responding on their behalf? Since about one in four of all adults are not self-respondents, measurement error can have potentially substantial effects.

The solution to this issue is not simple. The elimination of proxy reports would have an adverse impact on response rates. Travel surveys are already struggling with response rates due to the multistage interview process, the need to interview all household members, the burden on respondents due to the complexity of the questionnaire and request to maintain a diary, and the very short six-day data collection period. One change we plan to implement in future studies is to include an item identifying why a proxy respondent was needed, i.e., if there were language barriers, the person refused to participate, the person was away or was too ill to participate. This would perhaps improve our understanding of the mechanisms behind proxy usage and its impact on data quality.

Given some of the subgroup distinctions we have seen, e.g., differences in the factors predicting proxy response status between men and women, it might be more useful to pursue further examination of variations within subgroups.

This effort highlights the need for larger-scale experimental studies. We would also like to evaluate the quality of proxy responses provided by parents about their children. Finally, based on prior studies, it would seem that the relationship between the person providing the data and the person the data are being provided for might affect data quality.

References

Badoe, D. A. and Steuart, G. N (2002) Impact of Use of Message-Recording Devices, and Interviewing by Proxy on Large Scale Travel-Survey Conducted by Telephone. *Journal of Advanced Transportation*, Vol. 36 No. 1, pp 43-62.

Bickart, B., Blair, J. Sudman, S. and Menom, G. (1991) An Experimental Study of the Effects of Level of Participation on Proxy Report of Vacation Planning. *Proceedings from the Survey Research Methods Section*, American Statistical Association.

Collins M. A, Brick, J. M., Kim, K. and Stowe, P. (1997). *The Measurement of Participation in Adult Education*. NCES Technical Report 97-341, Washington, DC. U.S. Department of Education, National Center for Education Statistics downloaded 24 June 2004 from http://nces.ed.gov/pubs97/97341.pdf

Hess, J., Rothgeb, J. and Zukerberg, A. (1997) *Survey of Program Dynamics: Pretest Evaluation Report, Executive Summary*, Unpublished report. Bureau of the Census, downloaded 24 June 2004 from http://www.census.gov/srd/papers/pdf/sm97-10.pdf

Kojetin, B. A. and Miller, L. A. (1993) The Intrahousehold Communications Study: Estimating the Accuracy of Proxy Responses at the Dyadic Level. *Proceedings from the Survey Research Methods Section*, American Statistical Association.

Tables

Table 1: Percentage of adult proxy respondents in all households and households with more than one adult

Proxy Status	All Households			Households with More than One Adult		
	Sample Count	Unweighted	Weighted	Sample Count	Unweighted	Weighted
Proxy Report	10,196	22.3 %	22.5 %	10,174	26.0 %	26.0 %
Self Report	35,508	77.7 %	77.5 %	29,026	74.0 %	74.0 %
Total	45,704	100.0 %	100.0 %	39,200	100.0 %	100.0 %

NOTE: There is a slight change in the total number of proxy respondents since there were a few one-person households with a proxy respondent.
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person File

Table 2: Proxy presence by numbers of adults in household (in 100 percent households)

Number of Adults in Household	Percent of proxy respondents				
	0 proxies	1 proxy	2 proxies	3 proxies	4 proxies
1 adult	100 %	--	--	--	--
2 adults	49 %	50 %	--	--	--
3 adults	31 %	29 %	40 %	--	--
4 adults	19 %	16 %	20 %	45 %	--
5 adults	18 %	16 %	14 %	12 %	40 %

NOTE: 15 percent of six adult or more households have no proxy respondents.
NOTE: 100 percent households are those where all household adults have a completed interview.
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Household and Person Files

Table 3: Demographic characteristics of proxy respondents in households with more than one adult

Demographics	Weighted	
	Self	Proxy
Sex		
Male	45.5 %	62.0 %
Female	54.5 %	38.1 %
Age		
18-24	12.2 %	15.9 %
25-64	74.8 %	67.9 %
65 plus	13.1 %	16.2 %
Education		
High school or less	41.0 %	50.6 %
More than high school	59.0 %	49.4 %
Driver Status		
Driver	92.9 %	86.3 %
Non-driver	7.1 %	13.7 %
Commercial Driver		
Yes	77.6 %	79.9 %
No, non commercial driver	22.4 %	20.1 %
Worker Status†		
Employed	70.9 %	71.3 %
Non-employed	29.1 %	28.7 %
More than one job		
Yes	9.2 %	7.7 %
No, just one job	90.8 %	92.3 %
Country of Birth		
USA	87.4 %	85.1 %
Not USA	12.6 %	14.9 %
Medical Condition		
Yes	6.9 %	10.9 %
No	93.1 %	89.1 %
Diary Availability		
Kept Diary	64.0 %	66.9 %
Did Not Keep Diary	36.0 %	33.1 %
Travel Day (TD) Status		
Out of Town on TD	3.2 %	4.1 %
In Town on TD	96.8 %	95.9 %
Urbanucity†		
Urban	77.8 %	77.5 %
Rural	22.3 %	22.5 %
Total	100.0 %	100.0 %

† No significant difference in worker status and urbanicity at an alpha value of 0.05.
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Household and Person Files

Table 4: Key travel statistics by proxy status and household composition

Key Travel Measures	All Households			Households with More than Adult		
	Self	Proxy	Total	Self	Proxy	Total
Mean daily trips	4.5	3.7	4.3	4.5	3.7	4.3
Mean annual long distance trips	10.1	8.6	9.7	10.4	8.6	10.1
Mean weekly walk trips	4.2	3.1	4.0	4.2	3.1	3.9
Mean weekly bike trips	0.22	0.18	0.21	0.22	0.18	0.21
Public transit used on travel day	5.3 %	3.8 %	5.0 %	4.8 %	3.8 %	4.5 %
Public transit used in last two months	18.2 %	12.6 %	17.0 %	17.3 %	12.6 %	16.1 %

NOTE: All differences between proxy and self reports significantly different at an alpha value of 0.05.
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person, Daily Trip and Long Distance Trip Files

Table 5: Mean daily trips by purpose (in households with more than one adult)

Purpose	Self	Proxy
Work Trips	0.79	0.83
Work-related Trips	0.16	0.13
Family/Personal Trips	2.16	1.50
School/Church Trips†	0.21	0.21
Social/Recreation Trips	1.10	0.94
Other Trips†	0.03	0.03

† No significant difference between proxy and self trip means for school/church and other trips at an alpha value of 0.05.
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person and Daily Trip Files

Table 6: Demographic or 'true' factors related to total number key travel statistics (in households with more than one adult)

	Demographics	Daily Trips		Long Distance Trips		Walk Trips (Weekly)		Bike Trips (Weekly)	
		Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error
	Intercept	0.94	0.149*	0.51	0.098*	4.99	0.350*	0.56	0.065*
1	Travel-related Disability (1=yes)	-0.88	0.076*	-0.17	0.034*	-0.93	0.126*	-0.07	0.020*
2	Proxy Report (1=yes)	-0.65	0.034*	-0.15	0.027*	-1.15	0.089*	-0.07	0.013*
3	Employed (1=yes)	0.22	0.041*	0.22	0.025*	-0.60	0.107*	-0.00	0.015
4	Age (continuous)	-0.01	0.001*	0.00	0.001	-0.01	0.003*	-0.00	0.000*
5	Away from Home Travel Day (1=yes)	-1.16	0.107*	0.35	0.083*	1.06	0.282*	0.02	0.037
6	Non-driver (1-yes)	-1.40	0.054*	-0.33	0.027*	0.98	0.167*	0.16	0.036*
7	High School Education or Less (1=yes)	-0.52	0.036*	-0.21	0.029*	-0.40	0.091*	-0.02	0.013
8	Non-US Born (1=yes)	-0.33	0.053*	-0.26	0.031*	-0.82	0.104*	0.02	0.027
9	Male (1=yes)	-0.11	0.031*	0.29	0.022*	0.76	0.074*	0.15	0.014*

*Significant at an alpha value of 0.05
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person, Daily Trip and Long Distance Trip Files

Table 7: Predicting transit use for adults in households with more than one adult

	Demographics	Transit Used on Travel Day			Transit Used Last 2 Months		
		Parameter estimate	Standard error	Log Odds	Parameter estimate	Standard error	Log Odds
	Intercept	0.29	0.288		1.23	0.139*	
1	Travel-related Disability (1=yes)	0.12	0.141	1.13	-0.31	0.082*	0.74
2	Employed (1=yes)	0.18	0.092	1.20	0.19	0.045*	1.21
3	Age (continuous)	-0.02	0.002*	0.98	0.02	0.001*	0.98
6	Non-US Born (1=yes)	0.60	0.093*	1.82	0.58	0.056*	1.79
5	High School Education or Less (1=yes)	-0.46	0.079*	0.63	-0.75	0.046*	0.47
4	Non-driver (1-yes)	2.26	0.100*	9.58	1.95	0.057*	7.01
7	Male (1=yes)	0.15	0.067*	1.16	0.14	0.034*	1.09
8	Away from Home Travel Day (1=yes)	0.18	0.229	1.20	0.41	0.094*	1.51
9	Self (1=yes)	0.40	0.087*	1.49	0.49	0.043*	1.63

*Significant at an alpha value of 0.05
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person and Daily Trip Files

Table 8: Predicting a proxy response in households with more than one adult

	Demographics	Parameter estimate	Standard error	Log Odds
	Intercept	-0.75	0.131*	
1	Travel-related Disability (1=yes)	0.35	0.042*	1.42
2	Employed (1=yes)	0.14	0.036*	1.15
3	Age (continuous)	0.00	0.001	1.00
4	Non-US Born (1=yes)	0.09	0.049	1.10
5	High School Education or Less (1=yes)	0.30	0.027*	1.36
6	Non-driver (1-yes)	0.67	0.055*	1.96
7	Male (1=yes)	0.72	0.035*	2.05
8	Away from Home Travel Day (1=yes)	0.28	0.083*	1.33

*Significant at an alpha value of 0.05
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person File

Table 9: Predicting a proxy response for men and women (odds ratio) in households with more than one adult

	Demographics	Male	Female
1	Travel-related Disability (1=yes)	1.47	1.39
2	Employed (1=yes)	1.16	1.16
3	Age (continuous)*	0.998	1.00
4	Non-US Born (1=yes)*	0.79	1.65
5	High School Education or Less (1=yes)*	1.45	1.24
6	Non-driver (1-yes)*	1.64	2.10
7	Away from Home Travel Day (1=yes)	1.25	1.48

*Significant at an alpha value of 0.05
Source: 2001 National Household Travel Survey, U.S. Department of Transportation, Person File