Federal Highway Administration

Analysis of Partially Completed Households from the 1996 Bay Area Travel Study

Final Report

NuStats International
3006 Bee Caves Road, Ste A-300
Austin, Texas 78746

Voice 512-306-9065
Fax 512-306-9077

April 1998

EXECUTIVE SUMMARY

Most current data collection efforts for household travel surveys result in a pool of data from partially completed households. For the most part, this data is deemed unusable. Given the increase in the cost of data collection efforts, as well as the impact of sample non-response, there is a renewed interest in the value associated with this pool of largely ignored data.

This report reviewed the demographic data for 256 households with two or more members from the 1996 Bay Area Travel Study for which at least one member did not provide usable trip and activity data. The findings of the report show that the Bay Area Travel Study data without the inclusion of data for partially completed households is biased towards the following types of households:

- households with 4 or more persons,
- households with no workers or more than 2 workers,
- households with no vehicles,
- households earning less than $20,000,
- households earning at least $60,0000 but less than $75,000, and
- households that rent their residence.

In addition, at the person level, non-response bias was identified in the respondent age cohort of persons aged 19 to 29. When data from the partially completed households are added to the data set, the non-response bias was shown to be mitigated among the groups listed below. This suggests that more research is needed to identify methods for more consistent inclusion of these households into the travel models.

- households with 4 or more persons,
- households with more than 2 workers,
- households earning at least $60,000 but less than $75,000,
- households that rent their residence, and
- households with members ages 19 to 29.

The two main groups of non-respondents within a partially completed household are the secondary adults (adults other than the person recruited on the telephone) and teenagers. Further research and testing is needed to identify the appropriate mix of data collection techniques and respondent material design, and packaging that would encourage these household members to complete the surveys. This might include the redesign of respondent material packets to include individual envelopes within the larger household envelope that contain special appeals to the secondary household adults or the design of an abbreviated diary targeted towards teenagers.

Federal Highway Administration

Analysis of Partial Completed Households

Final Report

Table of Contents

INTRODUCTION 1
Report Purpose 1
Background 2
Definition of a Partially Completed Household 2
EVIDENCE OF NON-RESPONSE BIAS 3
IMPACT ON NON-RESPONSE BIAS OF PARTIALLY COMPLETED HOUSEHOLDS 7
PARTIALLY COMPLETED HOUSEHOLD CHARACTERISTICS..... 10
CONCLUSIONS AND RECOMMENDATIONS 13
Evidence of Non-Response..... 13
Impact on Non-Response Bias of Partially Completed Households 13
Characteristics of Partially Completed Households 14
Recommendations..... 14
Areas for Future Research 15

APPENDIX A: SUMMARY TABLES..... 17

Federal Highway Administration

Analysis of Partial Completed Households

Final Report

Tables

EVIDENCE OF NON-RESPONSE BIAS

Table 1: Comparison of Household Size with 1990 Census Data 3
Table 2: Comparison of Household Workers with 1990 Census Data 4
Table 3: Comparison of Household Vehicles with 1990 Census Data 4
Table 4 Comparison of Household Income with 1990 Census Data .. 4
Table 5 Comparison of Home Ownership with 1990 Census Data 5
Table 6 Comparison of Gender with 1990 Census Data 5
Table 7 Comparison of Household Member Ages with 1990 Census Data..... 5
Table 8 Comparison of Employment Status with 1990 Census Data. 6

IMPACT ON NON-RESPONSE BIAS OF PARTIALLY COMPLETED HOUSEHOLDS

Table 9 Combined Household Size versus 1990 Census Data..... 7
Table 10 Combined Household Workers versus 1990 Census Data ... 8
Table 11 Combined Household Vehicles versus 1990 Census Data ... 8
Table 12 Combined Household Income versus 1990 Census Data 8
Table 13 Combined Home Ownership versus 1990 Census Data 9
Table 14 Combined Household Member Age versus 1990 Census Data..... 9

PARTIALLY COMPLETED HOUSEHOLD CHARACTERISTICS

Table 15 Household Member Participation 10
Table 16 Likelihood of Participation by Gender..... 11
Table 17 Likelihood of Participation by Age 11
Table 18 Likelihood of Participation by Relationship..... 12
Table 19 Likelihood of Participation by Employment Status..... 12
Table 20 Likelihood of Participation by Student Status 12

CONCLUSIONS AND RECOMMENDATIONS

Table 21 Likelihood of Completion by Demographic Variable 14

APPENDIX A: SUMMARY TABLES

Table A1 Comparison of Household Size with 1990 Census Data - All Households..... 17
Table A2 Comparison of Household Workers with 1990 Census Data - All Households..... 17
Table A3 Comparison of Household Vehicles with 1990 Census Data - All Households..... 17
Table A4 Comparison of Household Income with 1990 Census Data - All Households..... 18
Table A5 Comparison of Household Ownership with 1990 Census Data - All Households..... 18

INTRODUCTION

The 1996 Bay Area Travel Study is comparable in design and standards to several household travel surveys conducted across the U.S. in the past decade. In particular, the contract required that the final data set be comprised only of households in which all members provided complete activity and travel behavior data. This requirement excluded from the data set those households in which not all of the household members provided data. These households are referred to as "partially completed" households.

The practice of excluding partially completed households from household travel survey data sets derives from the requirements of most transportation planning models. The calculation of a household trip rate is a base input into the models and requires data from all household members. While it is possible to apply certain imputation techniques so that partially completed households could approximate fully completed households, the transportation modeling community is not in agreement on the extent to which imputation should be used, if at all.

The focus of this paper is on the demographic data provided by the partially completed households, in order to determine if a systematic bias is being introduced into data sets by ignoring them. Specifically, this paper will present evidence of non-response bias in the survey sample and show the extent to which it can be mitigated by the inclusion of the partially completed households. In addition, characteristics of household members who tend not to participate will also be categorized.

Report Purpose

Most current data collection efforts for household travel surveys result in a pool of data from partially completed households. For the most part, this data is deemed unusable. Given the increase in the cost of data collection efforts, as well as the impact of sample non-response, there is a renewed interest in the value associated with this pool of largely ignored data.

This report will review the demographic data for 256 households with two or more members from the 1996 Bay Area Travel Study for which at least one member did not provide usable trip and activity data. The objectives of the report are as follows:

(1) to determine the extent to which non-response bias exists in the 1996 Bay Area Travel Study data set (which excludes partially completed households),

(2) to determine if the inclusion of partially completed households in the data set helps to reduce the non-response bias, and

(3) to determine what factors are associated with partially completed households so that survey methods can be improved.

Background

The 1996 Bay Area Travel Study is the fourth such study conducted in the region since 1965. The most recent study is very similar in content and design to the 1994 Oregon/SW Washington Household Activity and Travel Behavior Study. Both studies required all members of sampled households to record all travel and activities for a designated 48-hour period. In both studies, if any one member of a household did not participate, the household was deemed a partially completed household and therefore it was unusable in the sample. In the case of the Oregon study, the partially completed households were discarded. However, in the Bay Area study, the partially completed households were flagged as incomplete and processed using the same standards and processes for the fully completed households, under a contract from the Federal Highway Administration.

All data collection activities were performed using trained interviewers of NuStats International of Austin, Texas. Activity and travel behavior data were collected in two phases: January through May 1996, then September through December 1996. The result was the collection of demographic data for 5,861 recruited households, with 3,682 of those households providing complete activity and travel data for all household members. An additional 256 households provided activity and travel data for some but not all of the household members. These partially completed households form the basis of this analysis.

Definition of a Partially Completed Household

The group of non-respondents in the 1996 Bay Area Travel Study is comprised of three populations:

1. Households which did not participate at all (i.e., no demographic or travel data available).
2. Households which were recruited to participate but did not complete the travel diary (i.e., demographic data available but no travel data), and
3. Households which were recruited and some, but not all, members completed travel diaries (i.e., demographic data available for all and travel data for some, but not all, household members).

The third category comprises the “partially completed households” for purposes of this report, as defined below. For information on those who were recruited but did not provide any travel information at all, see Appendix A. **For purposes of this study, a partially completed household is defined as one in which at least one person did not provide complete activity and travel data**.

EVIDENCE OF NON-RESPONSE BIAS

Non-response bias in a survey data set occurs when certain individuals selected in a sample do not participate in the survey. The concern is that "*non-respondents will differ from respondents with regard to the survey variables, in which case the survey estimates based on the respondents alone will produce biased estimates of the overall population parameters*."[1] The purpose of this section is to examine the extent to which non-response bias is an issue in the 1996 Bay Area Travel Study sample.

The sampling plan for the 1996 Bay Area Travel Study was a random probability sample stratified by county, with an oversampling of residents that crossed the Bay Bridge. Telephone numbers were randomly generated and selected for inclusion in the study. All households in the nine-county region had an equal probability of selection.[2]

To ascertain the extent to which non-response bias is an issue with this data set, the household demographic characteristics for the 3,682 fully completed households are compared to population parameters - 1990 Census data. As shown in Tables 1 through 5, non-responding households in the Bay Area Travel Study appear to have the following characteristics:

- households with 4 or more persons,
- households with no workers or more than 2 workers,
- households with no vehicles,
- households earning less than $20,000,
- households earning between $60,000 but less than $75,000, and
- households that rent their residence.

Table 1
Comparison of Household Size with 1990 Census Data

Household Size	Fully Completed (n=3682)	Census Data	Difference
1 person	34.6%	26.0%	+8.6%
2 persons	34.1%	32.3%	+1.8%
3 persons	14.7%	16.7%	-2.0%
4+ persons	16.6%	25.0%	-8.4%
Total	**100%**	**100%**	
Avg. HH Size	**2.22**	**2.65**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all fully completed households).

[1] Kalton, Graham. Introduction to Survey Sampling. (Sage Publications: Beverly Hills, CA, 1987) p. 63.
[2] NuStats International. Technical Memorandum No. 5: Sample Design. Prepared for the Metropolitan Transportation Commission as part of the 1996 Bay Area Travel Study, March 1996.

Table 2
Comparison of Household Workers with 1990 Census Data

Household Workers	Fully Completed (n=3682)	Census Data	Difference
no workers	17.2%	20.9%	-3.7%
1 worker	41.9%	36.7%	+5.2%
2 workers	35.1%	33.1%	+2.0%
3+ workers	5.8%	9.3%	-3.5%
Total	**100%**	**100%**	
Avg. # workers	**1.32**	**1.40**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all fully completed households).

Table 3
Comparison of Household Vehicles with 1990 Census Data

Household Vehicles	Fully Completed (n=3682)	Census Data	Difference
no vehicles	5.7%	10.3%	-4.6%
1 vehicle	36.1%	32.1%	+4.0%
2 vehicles	38.3%	36.8%	+1.5%
3 vehicles	14.2%	14.5%	-0.3%
4+ vehicles	5.7%	6.3%	-0.6%
Total	**100%**	**100%**	
Avg. # vehicles	**1.81**	**1.74**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all fully completed households).

Table 4
Comparison of Household Income with 1990 Census Data

Household Income	Fully Completed (n=3162)	Census Data	Difference
Less than $10k	3.7%	9.3%	-5.6%
$10k but less than $20k	8.3%	11.6%	-3.3%
$20k but less than $40k	25.3%	13.2%	+12.1%
$40k but less than $60k	25.1%	26.6%	-1.5%
$60k but less than $75k	11.5%	33.4%	-21.9%
$75k but less than $100k	13.5%	9.9%	+3.6%
$100k or more	12.7%	9.1%	+3.6%
Total	**100%**	**100%**	
Avg. HH income	**$40-$45k**	**$50-$55k**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all fully completed households reporting income data).

Table 5
Comparison of Home Ownership with 1990 Census Data

Home Ownership	Fully Complete (n=3682)	Census Data	Difference
Own / buying	62.5%	56.4%	+6.1%
Rent	36.6%	43.6%	-7.0%
Other	0.9%	--	--
Total	**100%**	**100%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all fully completed households).

There are also important demographic differences in terms of household composition that potentially affect whether the household completes the survey task. The key person-level variables examined in comparison to 1990 Census Data were age, gender, and employment status. As shown in Tables 6 through 8 below, non-responding persons were primarily those between the ages of 19 to 29. There were no apparent differences with regard to gender or employment status.

Table 6
Comparison of Gender with 1990 Census Data

Gender	Fully Completed (n=8218)	Census Data	Difference
Male	49.3%	48.1%	+1.2%
Female	50.5%	51.9%	-1.4%
Refused	0.2%	--	--
Total	**100%**	**100%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all members of fully completed households).

Table 7
Comparison of Household Member Ages with 1990 Census Data

Age	Fully Completed (n=8055)	Census Data	Difference
Under 13	16.5%	17.3%	-0.8%
13 – 15	3.4%	3.4%	0.0%
16-18	2.9%	3.6%	-0.7%
19-29	12.8%	18.3%	-5.5%
30-39	18.0%	19.2%	-1.2%
40-49	18.3%	14.4%	3.9%
50-59	13.1%	8.9%	4.2%
60+	14.9%	14.9%	0.0%
Total	**100%**	**100%**	**100%**

Source: 1996 Bay Area Travel Study (Unweighted, Base=all members of fully completed households that reported age).

Table 8
Comparison of Employment Status with 1990 Census Data

Employment Status	Fully Completed (n=8054)	Census Data	Difference
Employed	70.1%	68.7%	+1.4%
Not employed	29.7%	31.3%	-1.6%
Employment status refused	0.2%	--	--
Total	**100%**	**100%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all members of fully completed households age 16 or older).

As shown above, non-response bias exists in the data set because particular types of households chose not to participate in the 1996 Bay Area Travel Study. If unaccounted for in the modeling process, it can impact the use of the data to the extent that the non-respondents differ from respondents on variables of interest to the users. The most affected variables will be the computed household activity and trip rates, since the characteristics identified are typically associated with households that make more trips than most.

IMPACT ON NON-RESPONSE BIAS OF PARTIALLY COMPLETED HOUSEHOLDS

When compared to 1990 census data, the 1996 Bay Area Travel Study shows areas of non-response by households that display the following characteristics:

- larger households with 4 or more persons,
- households with no workers or more than 2 workers,
- households with no vehicles,
- households earning less than $20,000, or $60,000 to $75,000,
- households that rent their residence, and
- households with members ages 19 to 29.

This means that the calculated household activity and trip rates may not be an accurate reflection of actual rates, since the above characteristics are typically associated with households that travel more frequently than others. An important research question is whether the demographics of the partially completed households would minimize the sample bias introduced by the non-responders. To answer that, the demographics for the partially completed households were combined with those for the fully completed households and again compared to census data. Any improvements (i.e., the collected data are closer to the census parameters) are noted in the tables below.

Table 9
Combined Household Size versus 1990 Census Data

Household Size	Full & Part Completed (n=3938)	Census Data	New Difference	Improved?
1 person	32.5%	26.0%	+6.5%	Yes
2 persons	34.6%	32.3%	+2.3%	No
3 persons	14.9%	16.7%	-1.8%	Yes
4+ persons	18.0%	25.0%	-7.0%	Yes
Total	**100%**	**100%**		
Avg. HH Size	**2.28**	**2.65**		

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households that provided any activity and travel data).

Table 10
Combined Household Workers versus 1990 Census Data

Household Workers	Full & Part Completed (n=3938)	Census Data	Difference	Improved?
no workers	16.6%	20.9%	-4.3%	No
1 worker	40.5%	36.7%	+3.8%	Yes
2 workers	36.9%	33.1%	+3.8%	No
3+ workers	6.0%	9.3%	-3.3%	Yes
Total	**100%**	**100%**		
Avg. # workers	**1.35**	**1.40**		

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households that provided any activity and travel data).

Table 11
Combined Household Vehicles versus 1990 Census Data

Household Vehicles	Full & Part Completed (n=3938)	Census Data	Difference	Improved?
no vehicles	5.5%	10.3%	-4.8%	No
1 vehicle	34.5%	32.1%	+2.4%	Yes
2 vehicles	39.4%	36.8%	+2.6%	No
3 vehicles	14.5%	14.5%	0.0%	Yes
4+ vehicles	6.0%	6.3%	-0.3%	Yes
Total	**100%**	**100%**		
Avg. # vehicles	**1.84**	**1.74**		

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households that provided any activity and travel data).

Table 12
Combined Household Income versus 1990 Census Data

Household Income	Full & Part Completed (n=3357)	Census Data	Difference	Improved?
Less than $10k	3.5%	9.3%	-5.8%	No
$10k but less than $20k	8.0%	11.6%	-3.6%	No
$20k but less than $40k	24.9%	13.2%	+11.7%	Yes
$40k but less than $60k	25.2%	26.6%	-1.4%	Yes
$60k but less than $75k	11.9%	33.4%	-21.5%	Yes
$75k but less than $100k	13.6%	9.9%	+3.7%	No
$100k or more	12.9%	9.1%	+3.8%	No
Total	**100%**	**100%**		
Avg. HH income	**$45-$50k**	**$50-$55k**		

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households that reported income data and any activity and travel data).

Table 13
Combined Home Ownership versus 1990 Census Data

Home Ownership	Full&Part Completed (n=3938)	Census Data	Difference	Improved?
Own / buying	63.0%	56.4%	+6.6%	No
Rent	36.1%	43.6%	-7.5%	No
Other	0.9%	--	--	
Total	**100%**	**100%**		

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households that provided activity and travel data).

Table 14
Combined Household Member Age versus Census Data

Age	Full & Part Completed (n=9019)	Census Data	Difference	Improved?
Under 13	16.3%	17.3%	-1.0%	No
13 – 15	3.7%	3.4%	0.3%	No
16-18	3.1%	3.6%	-0.5%	Yes
19-29	13.0%	18.3%	-5.3%	Yes
30-39	18.0%	19.2%	-1.2%	No
40-49	18.4%	14.4%	4.0%	No
50-59	13.1%	8.9%	4.2%	No
60+	14.4%	14.9%	0.5%	No
Total	**100%**	**100%**		

Source: 1996 Bay Area Travel Study (Unweighted, Base=all household members that reported age and any activity and travel data).

By including the partially completed households in the data set, non-response bias is mitigated among households with 4 or more persons. In addition, there is marginal (i.e., less than one percent) improvement among the following groups:

- households with more than 2 workers,
- households earning $60,000 to $75,000,
- households with members ages 19 to 29.

Any reduction in non-response bias means that the population parameters are more accurate and can be used with more confidence.

PARTIALLY COMPLETED HOUSEHOLD CHARACTERISTICS

As shown above, the inclusion of data from partially completed households results in a decrease of sample non-response for some specific population sub-groups. The next step in the research process is to determine if there are particular characteristics that distinguish these households so that data collection, sampling, or analytical methods can be prescribed to ensure that these types of households complete a travel survey. This is accomplished by examining the demographic composition of the household members that did participate in the survey as compared to those household members that did not. The chi-square statistic is used to determine if there are significant differences between responders versus non-respondents.

The data in Table 15 show the distribution of completed and partially completed households by household size. The fully completed households are shown in the cells where the number completing matches the household size (e.g., 92 percent of the 3-person households had all 3 members complete the study). The partials fill the other cells (e.g. 4.6 percent of all 3-person households had complete activity and travel data for only one household member). The distribution of households indicate that household size is negatively associated with fully completed households.

Table 15
Household Member Participation

Household Size	N	1	2	3	4	5	6	7	8	9	Total
1 person	1279	100.0%									**100.0%**
2 persons	1362	7.9%	92.1%								**100.0%**
3 persons	587	4.6%	3.2%	92.2%							**100.0%**
4 persons	463	8.4%	1.7%	3.2%	86.6%						**100.0%**
5 persons	172	9.3%	1.2%	1.7%	0.6%	87.2%					**100.0%**
6 persons	47	14.9%		2.1%		2.1%	80.9%				**100.0%**
7 persons	15	6.7%	6.7%	6.6%				80.0%			**100.0%**
8 persons	8								100.0%		**100.0%**
9 persons	5	20.0%								80.0%	**100.0%**

Source: 1996 Bay Area Travel Study (Unweighted, Base=3938 - all fully and partially completed households).

To identify those in the household that completed the survey as compared to those who did not, the following person-level variables were examined: gender, age, relationship, employment status, and student status. As shown in the tables below, the household members that are likely to complete the survey include those that were directly recruited (or the head of house), respondents between the ages of 40 and 59, and respondents that are not employed. Those household members not likely to complete the survey, rendering the household a partial -- and therefore unusable, include:

- respondents under the age of 30;
- respondents indirectly recruited (through the head of house), and
- respondents that are students.

Table 16
Likelihood of Participation by Gender

Gender	Completed Survey (n=338)	Didn't complete survey (n=463)	Statistically Different?
Male	49.1%	52.9%	No
Female	50.9%	46.0%	No
Refused	--	1.1%	--
Total	**57.8%**	**42.2%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all members of partially completed households). Statistically difference determined by results of chi-square analysis at 0.05 level of significance.

Table 17
Likelihood of Participation by Age

Age	Completed Survey (n=327)	Didn't complete survey (n=433)	Statistically Different?
Under 13	33.6%	66.4%	Yes
13 – 15	14.5%	85.5%	Yes
16-18	36.4%	63.6%	Yes
19-29	38.7%	61.3%	Yes
30-39	48.5%	51.5%	No
40-49	55.8%	44.2%	Yes
50-59	52.1%	47.9%	Yes
60+	41.2%	58.8%	No
Total	**43.0%**	**57.0%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all partially completed household members that reported age). Statistically difference determined by results of chi-square analysis at 0.05 level of significance.

Table 18
Likelihood of Participation by Relationship

Relationship	Completed Survey (n=338)	Didn't complete survey (n=463)	Statistically Different?
Head of House	95.7%	4.3%	Yes
Spouse	10.2%	89.8%	Yes
Child	25.8%	74.2%	Yes
Extended Family (related)	13.3%	86.7%	Yes
Non-related	6.7%	93.3%	Yes
Total	**42.2%**	**57.8%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all partially completed household members). Statistically difference determined by results of chi-square analysis at 0.05 level of significance.

Table 19
Likelihood of Participation by Employment Status

Employment Status	Completed Survey (n=294)	Didn't complete survey (n=340)	Statistically Different?
Employed	50.8%	49.2%	No
Not employed	58.8%	41.2%	Yes
Total	**53.6%**	**46.4%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all partially completed household members age 16 or older). Statistically difference determined by results of chi-square analysis at 0.05 level of significance.

Table 20
Likelihood of Completion Status by Student Status

Student Status	Completed Survey (n=338)	Didn't complete survey (n=463)	Statistically Different?
Student	32.7%	67.3%	Yes
Not a Student	46.7%	53.3%	No
Total	**53.6%**	**46.4%**	

Source: 1996 Bay Area Travel Study (Unweighted, Base=all partially completed household members). Statistically difference determined by results of chi-square analysis at 0.05 level of significance.

CONCLUSIONS AND RECOMMENDATIONS

The purpose of this paper was to review the demographic data for 256 households that participated in the 1996 Bay Area Travel Study, in which at least one household member did not provide usable trip and activity data. These households are typically referred to as partially completed households and, in most cases, are excluded from the transportation planning process. Given the increase in data collection costs, as well as the resulting impact on non-response for the data set as a whole, there is a renewed interest in the value associated with this pool of largely ignored data.

Evidence of Non-Response

Non-response in a survey is characterized by non-participation from certain sampled individuals. The concern is that biased estimates for overall survey population parameters could result. To determine the extent to which non-response bias existed in the 1996 Bay Area Travel Study data set, the household demographics for the fully completed households were compared to 1990. Non-response was identified in the following respondent groups:

- households with 4 or more persons,
- households with no workers or more than 2 workers,
- households with no vehicles,
- households earning less than $20,000,
- households earning at least $60,0000 but less than $75,000, and
- households that rent their residence.

In addition, at the person level, non-response bias was identified in the respondent age cohort of persons aged 19 to 29.

Impact on Non-Response Bias of Partially Completed Households

The second research question for this paper was to determine if the inclusion of partially completed households into the data set resulted in a reduction of non-response. To answer this question, demographic data from the partially completed households were included in the data set and again compared to 1990 census data. By doing so, there was a reduction in the non-response by the following groups:

- households with 4 or more persons,
- households with more than 2 workers,
- households earning at least $60,000 but less than $75,000,
- households that rent their residence, and
- households with members ages 19 to 29.

Characteristics of Partially Completed Households

The final question investigated as part of this report was to determine what factors were associated with partially completed households. By examining the demographic characteristics of responders versus non-responders within the 256 partially completed households, the responders were found to be statistically different from the non-responders for the following groups:

Table 21
Likelihood of Completion by Demographic Variable

Demographic Characteristic	Completed Survey	Didn't complete survey	Statistically Different?
Relationship	Head of House	All others	Yes
Age	40 to 59	under 30	Yes
Employment Status	Not employed	(no difference)	Yes
Student Status	(no difference)	Student	Yes

Source: 1996 Bay Area Travel Study (Unweighted, Base=all household members). Statistically difference determined by results of chi-square analysis at 0.05 level of significance.

Gender was not found to have an impact completion status of any household members at the 0.05 confidence interval, using the chi-square measure of significance.

Recommendations

Several data collection and sampling techniques can be used to increase the likelihood of participation among households and persons that represent partially completed household characteristics. In addition, study pre-tests offer an opportunity to study the affects of non-response bias and modify sampling and data collection methods planned for the full study.

Specifically, any pre-test data collection instructions should require the collection of activity and travel data from any respondents that completed diary, regardless of how many other household members did the same. This initial picture of non-response will help to identify potential non-response bias in the full study and modifications can be made to the sampling plan (i.e., oversampling) or data collection instructions to help minimize non-response from targeted households. The entry and processing of this data need not take place, but the documentation of who in the household provided data will help to characterize the non-response that can be expected in the actual data collection effort.

Once the likely non-responding sample members are identified, it is be possible to modify sampling plans to oversample households meeting specific characteristics, in an effort to have proper representation in the full data set. An alternative to oversampling would be to provide special treatment to households exhibiting these characteristics for the remaining phases of the study. While not exhaustive, a list of possible

special treatment efforts includes the following:

- the CATI script can be modified to identify specific household characteristics and special scripts can be introduced for those households,
- the packet of respondent materials can be modified to include more information about the importance of the study or specific appeals for participation (for the indirectly recruited household members),
- the packet of respondent materials can include a person-level incentive,
- the packet of respondent materials can include a postage-paid envelope to make it easier for everyone in the household to complete their diaries and mail them back,
- the retrieval survey can include a question to gather information on who else from the household accompanied the respondent on any particular trip,
- the retrieval sample can be marked for priority handling and routed to interviewers trained in securing participation by all household members, and
- partially completed households could be tracked and, if not firm refusals, rescheduled for travel days in which all members could participate.

Areas for Future Research

While the research conducted in the course of preparing this paper have provided evidence as to the composition of partially completed households and who within those households tends to not participate, there are several questions still outstanding.

The first question is how to integrate these households into the data sets used for transportation planning modeling. The whole issue of data imputation is extremely volatile and there is no agreement in the field as to when to include imputed data, if at all. There are some excellent sources of information on various imputation techniques[3] and current case studies of application in the modeling process, particularly in previous Bay Area Travel Studies and the recent Dallas Travel Study. However, the field is lacking recommended treatments as well as the associated pros and cons for using the data.

[3] see Table 2 in David Hensher's Issues in the Pre-Analysis of Panel Data as published in the Transportation Research A, Vol. 21A, No. 4/5, pp. 265-285, 1987.

A second issue needing research is the question of person-based sampling. The reason the partially completed households exist is because current travel studies are household-based. This is largely due to the need to calculate a household trip rate using the resultant data sets as well as the household interactions provided by the travel data. In some specific cases, data for head of house or commuters may be investigated at the person level. A movement towards person-based sampling would greatly reduce the non-response bias in the data sets. However, the integration of this into the modeling process needs to be investigated, as well as methods for capturing the interactions with other household members.

The funding sources for large scale investigations into these issues are scarce. However, the limited research dollars can be leveraged by adding methodological options to pre-tests or data collection tasks in current and forthcoming travel studies. Not only would this be more cost effective, but it would also provide insight into the viability of optional methods for addressing the issues of reducing and more effectively dealing with non-response.

APPENDIX A: SUMMARY TABLES

Table A1
Comparison of Household Size with 1990 Census Data - All Households

Household Size	Recr but no Travel Data (n=1923)	Partial Complete (n=256)	Fully Complete (n=3682)	Full & Part Completes (n=3938)	Census Data
1 person	19.2%	0.0%	34.6%	32.5%	26.0%
2 persons	29.8%	44.2%	34.1%	34.6%	32.3%
3 persons	19.5%	18.0%	14.7%	14.9%	16.7%
4+ persons	31.5%	37.8%	16.6%	18.0%	25.0%
Total	**100%**	**100%**	**100%**	**100%**	**100%**
Avg. HH Size	**2.84**	**3.13**	**2.22**	**2.28**	**2.65**

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households).

Table A2
Comparison of Household Workers with 1990 Census Data - All Households

Household Workers	Recr but no Travel Data (n=1923)	Partial Complete (n=256)	Fully Complete (n=3682)	Full & Part Completes (n=3938)	Census Data
no workers	15.4%	7.8%	17.2%	16.6%	20.9%
1 worker	36.1%	19.5%	41.9%	40.5%	36.7%
2 workers	35.6%	62.5%	35.1%	36.9%	33.1%
3+ workers	12.9%	10.2%	5.8%	6.0%	9.3%
Total	**100%**	**100%**	**100%**	**100%**	**100%**
Avg. # workers	**1.51**	**1.82**	**1.32**	**1.35**	**1.40**

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households).

Table A3
Comparison of Household Vehicles with 1990 Census Data - All Households

Household Vehicles	Recr but no Travel Data (n=1923)	Partial Complete (n=256)	Fully Complete (n=3682)	Full & Part Completes (n=3938)	Census Data
no vehicles	8.0%	2.7%	5.7%	5.5%	10.3%
1 vehicle	28.2%	12.1%	36.1%	34.5%	32.1%
2 vehicles	38.3%	55.1%	38.3%	39.4%	36.8%
3 vehicles	16.4%	19.1%	14.2%	14.5%	14.5%
4+ vehicles	9.1%	10.9%	5.7%	6.0%	6.3%
Total	**100%**	**100%**	**100%**	**100%**	**100%**
Avg. # vehicles	**1.97**	**2.30**	**1.81**	**1.84**	**1.74**

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households).

Table A4
Comparison of Household Income with 1990 Census Data - All Households

Household Income	Recr but no Travel Data (n=1439)	Partial Complete (n=195)	Fully Complete (n=3162)	Full & Part Completes (n=3357)	Census Data
Less than $10k	4.1%	0.0%	3.7%	3.5%	9.3%
$10k but less than $20k	9.8%	3.0%	8.3%	8.0%	11.6%
$20k but less than $40k	27.4%	20.6%	25.3%	24.9%	13.2%
$40k but less than $60k	27.0%	27.7%	25.1%	25.2%	26.6%
$60k but less than $75k	10.0%	17.4%	11.5%	11.9%	33.4%
$75k but less than $100k	10.6%	15.9%	13.5%	13.6%	9.9%
$100k or more	11.1%	15.4%	12.7%	12.9%	9.1%
Total	**100%**	**100%**	**100%**	**100%**	**100%**
Avg. HH income	**$45-$50k**	**$50-$60k**	**$40-$45k**	**$45-$50k**	**$50-$55k**

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households reporting income data).

Table A5
Comparison of Household Ownership with 1990 Census Data - All Households

Household Ownership	Recr but no Travel Data (n=1439)	Partial Complete (n=195)	Fully Complete (n=3162)	Full & Part Completes (n=3357)	Census Data
Own / buying	55.0%	69.1%	62.5%	63.0%	56.4%
Rent	43.3%	28.9%	36.6%	36.1%	43.6%
Other	1.7%	2.0%	0.9%	0.9%	--
Total	**100%**	**100%**	**100%**	**100%**	**100%**

Source: 1996 Bay Area Travel Study (Unweighted, Base=all households).