Do Characteristics of RDD Survey Respondents Differ According to Difficulty of Obtaining Response?

Pheny Z. Weidman
U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics, 1200 New Jersey Avenue, SE Washington, DC 20590

Abstract
Unit nonresponse in any survey may introduce substantial bias or error into survey estimates when (i) the number of nonrespondents is large relative to the sample size or (ii) the characteristics of nonrespondents differ greatly from those of respondents. Investigators of many well-established surveys expend considerable effort and money to recruit difficult-to-reach respondents. This additional effort and expense are only worthwhile if they result in a sample that is representative of the target population. In this study, logistic regression models are developed to compare characteristics of willing, accessible respondents with those of their less accessible or less willing counterparts to determine whether or not the two groups differ with respect to demographic characteristics. Data collected from the Bureau of Transportation Statistics' Omnibus Household Survey (OHS), a transportation customer satisfaction survey using a random digit dialing method, will be analyzed. In addition, because the sample size of the annual OHS is very small, only slightly more than 1,000 individuals, multiple years of data will be used to augment sample size for this study.

Key words: RDD survey, Characteristics of respondents, Selection bias, Unit nonresponse bias

1. Introduction

In describing the virtues of the sample survey, Sidney Verba has written that “surveys produce just what democracy is supposed to produce—equal representation of all citizens. The sample survey is rigorously egalitarian; it is designed so that each citizen has an equal chance to participate and an equal voice when participating” (Verba 1996, p. 3). Although Verba acknowledges that the people interviewed in surveys are not truly random samples, he sees surveys as much closer to the egalitarian ideal than any other venue from which citizens can be heard (Keeter et al., 2000; Verba, 1996).

However, unit nonresponse in a household survey may introduce substantial bias into estimates when (i) the number of nonrespondents is large relative to the sample size or (ii) the characteristics of nonrespondents differ greatly from those of respondents (Groves, 2006; Voigt et al., 2003; Lin and Schaeffer, 1995). Unit nonresponse occurs when a sampled unit such as a person, household, or organization fails to respond to a survey. Many household surveys expend considerable effort and money to interview difficult-to-contact respondents so that participants will be representative of the population of interest. Although the response rate generally improves as effort to recruit

reluctant respondents increases, such effort typically is both expensive and time-consuming. It is only worthwhile if it results in a set of respondents that is representative of the target population.

Several studies have compared (a) respondents who initially refuse and later agree to participate with respondents who readily agree to participate or (b) early respondents to late respondents. Many have found that reluctant or difficult-to-contact respondents are older and less educated than amenable respondents. However, research findings with respect to income, occupation, race, and marital status have been inconsistent (Johnson et al., 2006; Voigt et al., 2003; Etter and Perneger, 1997; Triplett et al., 1996; Kaldenberg et al., 1994; Kristal et al., 1993; Lavrakas et al., 1992; Holt et al., 1991; Groves, 1989; Fitzgerald and Fuller, 1982).

This study was undertaken to address the question of whether differences exist in demographic characteristics between random digit dialing (RDD) survey respondents according to the level of effort required to recruit them.

2. Data and Method

The Bureau of Transportation Statistics' (BTS) Omnibus Household Survey (OHS) was used in this study. The OHS is a Customer Satisfaction Survey using a list-assisted RDD methodology. It assesses the general public's perception of, expectations from, and satisfaction with the nation's transportation system by interviewing persons in randomly selected telephone households. It is conducted by the BTS with joint funding from the Transportation Security Administration (TSA). The survey was first conducted in 2000. The analysis this study reports on was based on OHS data collected during Nov. 2006 and Nov. 2007.

2.1 Sample Design

The target population for the OHS is the noninstitutionalized population aged 18 years or older who currently live in the United States. To ensure that a sample of telephone numbers is geographically representative, telephone prefixes are stratified by their associated Census Bureau divisions and metropolitan status (Bureau of Transportation Statistics, 2007; 2006). The rate at which households are sampled for the OHS is the same across these strata. In the last stage of sample selection, one randomly selected person age 18 or older in each sampled household is designated for participation in the OHS. The list-assisted random digit dialing methodology was employed to generate the desired sample. List-assisted refers to the use of commercial lists of directory-listed telephone numbers to increase the likelihood of dialing household residences. This method gives unlisted telephone numbers the same chance to be selected as directory-listed numbers.

In both 2006 and 2007, the OHS interviews were conducted and completed during the month of November. However, interviews were not conducted on Thanksgiving Day in either year.

If a selected person could not be contacted or was not available, an interviewer would try to call back later at a different time. Although the maximum number of calls back to a non-contact person was set as 60, the highest number recorded in our data was 80. The disposition of each call was recorded by the interviewers.

The OHS final analysis weight is the product of five components: 1) base sampling weight**;** 2) adjustment for unit nonresponse; 3) adjustment for households with multiple telephone numbers; 4) adjustment for selecting an adult within a sampled household; and 5) post-stratification adjustments to the target population.

2.2 Methods

Because the respondent sample size of the annual OHS is very small – only slightly more than 1,000 individuals – multiple years of data were needed to augment sample size for this study. However, this study was able to use only the 2006 and 2007 OHSs, because interview history and disposition information associated with other years of OHS interviews conducted by the same telephone survey firm were not available. To examine potential differences with respect to age, gender, race, ethnicity, household income, education level, and metropolitan status of the OHS respondents according to the difficulty of obtaining response, first the interview history and disposition data of the 2006 and 2007 OHSs were merged with their respective OHS data files. It was discovered that interview history information was missing for 192 individuals in the 2006 OHS data file. Subsequently, those people were excluded from this study. Indicators of the level of difficulty of eliciting a response based on each sample unit's calling history were then constructed for the remaining combined 903 and 1,016 sample units from the 2006 and 2007 OHSs, respectively. The final analysis weight of each sample unit in the combined data file was calculated as half of its individual year final weight, so that the total final weight is the average of the 2006 and 2007 target populations. Each OHS sample unit was classified into one of three categories: early respondents, difficult-to-contact respondents, and initial refusers. Individuals who completed telephone interviews by no later than the 3rd phone call were classified as early respondents. Those who completed telephone interviews only after at least four phone calls were classified as difficult-to-contact respondents. Initial refusers were those respondents who initially refused regardless of how many phone calls were made before the completion of interview. Such categorization of respondents was actuated by the assumption that the characteristics of people who are difficult to interview become more like those of nonrespondents as the difficulty of interviewing them increases -- a common assumption that is widely held both implicitly and explicitly by many studies in the statistical literature regarding methods of adjusting for survey nonresponse bias (Johnson et al., 2006; Voigt et al., 2003; Etter and Perneger, 1997; Triplett et al., 1996; Lin and Schaeffer, 1995; Kristal et al., 1993; Potthoff et al., 1993; Lavrakas et al., 1992; Smith, 1984; Thomsen and Siring, 1983; Fitzgerald and Fuller, 1982; Stinchcombe et al., 1981; Filion, 1976).

All variables were included simultaneously in the multinomial logistic regression models with the early respondents used as the reference group, using SUDAAN MULTILOG Procedure (RTI International, 2009). After adjusting for all other variables analyzed, *p* values were computed for individual variables using Wald's chi-square test. Any difference was considered statistically significant if its corresponding *p* value was less than 0.10. Results are presented in the tables.

3. Results and Discussion

3.1 Results

3.1.1 Multinomial Models

Table 1 summarizes the results of main effect tests derived from the initial multinomial logistic regression model with variable difficulty of obtaining a response as the dependent variable and age, gender, race, ethnicity, household income, education level, and metropolitan status as predictors or covariates. The Wald's chi-square test was used to evaluate these effects. In comparison of the difficult-to-contact respondents and initial refusers with early respondents, initial analysis indicated that ethnicity (p = .04), age (p = .02), education (p = .02), and income (p = .07) were statistically significant covariates.

Table 1: Summary of the Main Effects Tests Conducted by Generalized Logit Model for the Level of Difficulty to Obtain Response Study (Including All Seven Predictors and Using Unrestricted Data Set)

Treatment/Contrast	Degree of freedom	Wald's χ^2	p-value
Overall Model	34	274.84	0.000
Model - Intercept	32	70.31	0.000
Metropolitan Status	2	0.64	0.727
Gender	2	0.00	0.998
Ethnicity	2	6.65	0.05*
Race	2	0.83	0.660
Age	10	21.71	<0.05*
Education	8	18.02	<0.05*
Income	6	11.83	0.066*

* $p < 0.10$ compared with early responders. Data were adjusted for all other variables.

Total Number of Observations in Data File = 1,919

Observations used in the analysis = 1,376

Initial Refuser n = 147

Difficult-to-Contact n = 620

Early Responder n = 609

SOURCE: U.S. Department of Transportation, Bureau of Transportation Statistics, Omnibus Household Survey, November 2006 and 2007.

Analysis information indicated that there were 1,919 individuals on the file and that only 1,376 were used in the analysis. 543 individuals were deleted due to missing values on one or more predictors. The type of information collected by the OHS made not only any data imputation difficult but it also raised serious doubts on whether imputed data could actually improve the precision of resulting estimates. Thus, it was decided that no imputation for missing data would be used in this study. Among the seven predictors, income was primarily responsible for the severe sample loss in the analysis. It contained the highest percentage of missing value, approximately 26%. In order to determine its impact on the significance of the other predictors, a second multinomial logistic regression model was fitted with the same dependent variable and set of predictors except for the income variable.

Table 2 summarizes the results of main effect tests derived from the second multinomial logistic regression model. It indicated that without income as a covariate only age (p = .01) and education (p = .10) were statistically significant. The main effect of ethnicity was no longer statistically significant (p = .22). Analysis information indicated that there were 1,919 individuals on the file and that 1,792 were used in the analysis. 127 individuals were deleted due to missing values on one or more predictors. Therefore, the difference between first and the second model was more than exclusion of income covariate alone. In addition to the absence of income variable, the second model used 416 more sample persons for its estimation than the first model. Hence, any difference in their

Table 2: Summary of the Main Effects Tests Conducted by Generalized Logit Model for the Level of Difficulty to Obtain Response Study (with Only Six Predictors-without Income Variable, and Using Unrestricted Data Set)

Treatment/Contrast	Degree of freedom	Wald's χ^2	p-value
Overall Model	28	310.55	0.000
Model - Intercept	26	62.14	0.000
Metropolitan Status	2	3.45	0.179
Gender	2	0.39	0.822
Ethnicity	2	3.06	0.2168
Race	2	3.78	0.151
Age	10	24.86	<0.01*
Education	8	13.34	0.10*

* $p < 0.10$ compared with early responders. Data were adjusted for all other variables.
Total Number of Observations in Data File = 1,919
Observations used in the analysis = 1,792
Initial Refuser n = 212
Difficult-to-Contact n = 786
Early Responder n = 794
SOURCE: U.S. Department of Transportation, Bureau of Transportation Statistics, Omnibus Household Survey, November 2006 and 2007.

respective analytical results might be due to the effects of (1) exclusion of income, (2) increased sample size, (3) combined effect of exclusion of income and increased sample size, and (4) simply by chance alone. To attempt to isolate the effect of excluding income from the model alone, a restricted data file was created which contained only individuals who had no missing value on any of the seven predictors age, gender, race, ethnicity, household income, education level, and metropolitan status. The previously discussed second multinomial logistic regression model was refitted with the restricted data. Table 3 summarizes the results of main effect tests derived from this model.

Analysis indicated that ethnicity (p = .04), age (p = .01), and education (p = .03) were statistically significant covariates. In the models without income for the two data sets, as well as for the model including income, both age and education were statistically significant. So it could be concluded that they were truly statistically significant treatment effects. However, this was not the case with ethnicity, as it was not significant in the model without income for the larger data set. Thus the analyses were inconclusive about the significance of ethnicity.

Table 3: Summary of the Main Effects Tests Conducted by Generalized Logit Model for the Level of Difficulty to Obtain Response Study (with Only Six Predictors-without Income Variable, and Using Restricted Data Set)

Treatment/Contrast	Degree of freedom	Wald's χ^2	*p*-value
Overall Model	28	265.29	0.000
Model - Intercept	26	58.66	0.000
Metropolitan Status	2	0.93	0.630
Gender	2	0.02	0.990
Ethnicity	2	6.50	<0.05*
Race	2	0.31	0.857
Age	10	23.30	<0.01*
Education	8	17.20	<0.05*

* $p < 0.10$ compared with early responders. Data were adjusted for all other variables.

Total Number of Observations in Data File = 1,376

Observations used in the analysis = 1,376

Initial Refuser n = 147

Difficult-to-Contact n = 620

Early Responder n = 609

SOURCE: U.S. Department of Transportation, Bureau of Transportation Statistics, Omnibus Household Survey, November 2006 and 2007.

3.1.2 Comparison of Difficult-to-Contact Responders and Initial Refusers with Early Responders

The difficult-to-contact responders were more likely to be younger and live in a family with a moderate household income (at least $30,000 and < $50,000 annual household income) than early responders (table 4). Initial refusers were more likely to have attended some college and less likely to have low household income (less than $30,000) than early responders.

Table 4: Demographic Characteristics and Metropolitan Status (%) of Early Responders, Difficult-to-Contact Responders, and Initial Refusers from the Omnibus Household Survey, November 2006 and 2007

	Early Responders % (n = 609)	Difficult-to-Contact Responders % (n = 620)	Initial Refusers % (n = 147)
Metropolitan Status			
MSA Area	76.85	79.05	80.87
p value		0.92	0.43
Non-MSA Area	23.15	20.95	19.03

Table 4: Demographic Characteristics and Metropolitan Status (%) of Early Responders, Difficult-to-Contact Responders, and Initial Refusers from the Omnibus Household Survey, November 2006 and 2007--continued

	Early Responders % (n = 609)	Difficult-to-Contact Responders % (n = 620)	Initial Refusers % (n = 147)
Race			
White	77.82	73.06	71.50
p value		0.47	0.43
Non-White	22.18	26.94	28.50
Ethnicity			
Hispanic	5.32	11.84	11.71
p value		<0.05*	0.13
Non-Hispanic	94.68	88.16	88.29
Gender			
Male	49.53	48.78	50.16
p value		0.95	0.99
Female	50.47	51.22	49.84
Age (years) at reference date			
18-34	22.96	30.03	21.85
p value		0.08*	0.82
35-44	17.39	23.99	21.83
p value		<0.05*	0.60
45--54	22.60	20.81	22.42
p value		0.48	0.87
55-64	17.37	13.69	15.06
p value		0.98	0.78
65-74	11.67	6.00	11.87
p value		0.16	0.91
≥ 75	8.00	5.47	6.97
Education			
Up to High School Graduate	38.15	35.88	39.65
p value		0.76	0.14
Some College	15.46	17.30	22.40
p value		0.87	<0.05*
AA Degree	16.23	11.73	17.79
p value		0.11	0.21
BA or BS Degree	16.88	21.83	11.19
p value		0.37	0.97
Graduate Degree	13.29	13.25	8.98

Table 4: Demographic Characteristics and Metropolitan Status (%) of Early Responders, Difficult-to-Contact Responders, and Initial Refusers from the Omnibus Household Survey, November 2006 and 2007--continued

	Early Responders % (n = 609)	Difficult-to-Contact Responders % (n = 620)	Initial Refusers % (n = 147)
Income			
< $30,000	28.37	21.25	22.37
p value		0.51	0.07*
$30,000 to < $50,000	18.12	23.61	25.17
p value		0.08*	0.98
$50,000 to < $100,000	35.31	38.32	31.48
p value		0.45	0.32
$100,000 or more	18.21	16.82	20.98

* $p < 0.10$ compared with early responders. Data were adjusted for all other variables.
Total Number of Observations in Data File = 1,919

SOURCE: U.S. Department of Transportation, Bureau of Transportation Statistics, Omnibus Household Survey, November 2006 and 2007.

3.2 Discussion

This study examines whether differences exist in demographic characteristics between RDD survey respondents according to the level of effort required to recruit them. Although it found several significant differences when difficult-to-contact responders and initial refusers were compared with early responders, it did not find any clear pattern of characteristics that could be useful in generalizing these differences. This conclusion is consistent with those from similar studies in the literature (Voigt et al., 2003; Triplett et al., 1996; Lin and Schaeffer, 1995; Kristal et al., 1993; Lavrakas et al., 1992; Fitzgerald and Fuller, 1982). The differences in reluctant respondent characteristics, in similarly designed studies, might be due to variations over time or reflect geographical differences or variations in data collection mode. However, findings of this study suggest that additional effort expended in recruiting reluctant respondents by surveys such as the OHS would most likely result in more accurate estimates of population characteristics that are of interest in survey research.

There are certain limitations associated with this study. Its ability to find differences between the response groups might be limited by the relatively small sample size of the three comparison groups, especially the initial refusers group. In addition, for this particular study, the missing data rate was very high. About 28.3% of the cases didn't have information on one or more predictors. This is a common problem encountered by household surveys when income information is collected. Recent study shows that the typical item nonresponse rate to income question is between 20% and 40% (Yan et al., 2010)

Despite its limitations, this study provides concrete evidence to justify the additional effort and expense used in recruiting reluctant respondents by the OHS. Although findings of the study support the survey's current recruitment strategy for reluctant

responders, it is recommended that more studies with larger samples are needed. The author suggests that future studies on reluctant responder characteristics might consider building differences in data collection mode and geographic area as additional comparison factors to be studied.

References

Bureau of Transportation Statistics (2007), “Survey Documentation for the Bureau of Transportation Statistics Omnibus Survey Program (Public Use), November 2007,” Washington, D.C.: Bureau of the Transportation Statistics.

Bureau of Transportation Statistics (2006), “Survey Documentation for the Bureau of Transportation Statistics Omnibus Survey Program (Public Use), November 2006,” Washington, D.C.: Bureau of the Transportation Statistics.

Cochran, W. G. (1983), "Historical Perpective," In Incomplete Data in Sample Surveys. Vol. 2, Theory and Bibliographies, ed. William G. Madow, Ingram Olkin, and Donald B. Rubin, pp. 11-25. New York: Academic Press.

Drew, J. H. and Fuller, W. A. (1980), “Modeling Nonresponse in Surveys with Callbacks,” Proceedings of Survey Research Methods Section, American Statistical Association pp. 639-42.

Etter, J. and Perneger, T. V. (1997), “Analysis of Nonresponse Bias in a Mailed Health Survey,” Journal of Clinical Epidemiology 50:1123-28.

Filion, F. L. (1976), “Exploring and Correcting Nonresponse Using Follow-Ups of Nonrespondents,” Pacific Sociological Review 19:401-08.

Fitzgerald, R. and Fuller, L. (1982), “I Hear You Knocking But You Can’t Come in: The Effect of Reluctant Respondents and Refusers on Sample Survey Estimates,” Sociological Methods and Research 11:3-32.

Groves, R. M. (2006), “Nonresponse Rate and Nonresponse Bias in Household Surveys,” Public Opinion Quarterly 70:646-75.

Groves, R. M. (1989), Survey Errors and Survey Costs. Probing the Causes of Nonresponses and Efforts to Reduce Nonresponse. New York: John Wiley and Sons, Inc, 1989:185.

Holt, V. L., Daling, J. R., Stergachis, A., et al. (1991), "Results and Effect of Refusal Recontact in a Case-Control Study of Ectopic Pregnancy," Epidemiology 2:375-9.

Johnson, T. P., Cho, Y. I., Campbell, R. T., and Holbrook, A. L. (2006), “Using Community-Level Correlates to Evaluate Nonresponse Effects in a Telephone Survey,” Public Opinion Quarterly 70:704-19.

Kaldenberg, D., Koenig, H. F., Becker, B. W. (1994), "Mail Survey Response Rate Patterns in a Population of the Elderly," Public Opinion Quarterly 58:68-76.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., and Presser, S. (2000), “Consequences of Reducing Nonresponse in a National Telephone Survey,” Public Opinion Quarterly 64:125-148.

Kristal, A. R., White, E., Davis, J. R. et al. (1993), “Effects of Enhanced Calling Efforts on Response Rates, Estimates of Health Behavior, and Costs in a Telephone Health Survey Using Random-digit Dialing,” Public Health Rep 108:372-9.

Lavrakas, P. J., Bauman, S. L., and Merkle, D. M. (1992), “Refusal Report Forms (RRF), Refusal Conversions, and Non-response Bias,” Presented at the 47th Annual Meeting of the American Association for Public Opinion Research, St. Petersburg, Florida, May 15-19.

Lin, I. F. and Schaefer, N. C. (1995), “Using Survey Participants to Estimate the Impact of Nonparticipation,” Public Opinion Quarterly, 59:236-58.

Little, R. J. A., and Rubin, D. B. (1987), Statistical Analysis with Missing Data. New York: Wiley.

Rubin, D. B. (1987), Multiple Imputation for Nonresponse in Surveys. New York: Wiley.

Potthoff, R., Manton, K., and Woodbury, M. (1993), "Correcting for Nonavailability Bias in Surveys by Weighting Based on Number of Callbacks," Journal of the American Statistical Association, 88:1197-207.

Smith, T. W. (1984), "Estimating Nonresponse Bias with Temporary Refusals," Sociological Perspectives, 27(4): 473-89.

Stinchcombe, A. L., Jones, C., and Sheatsley, P. (1981), "Nonresponse Bias for Attitude Questions," Public Opinion Quarterly, 45(3): 359-75.

RTI International. (2009), SUDAAN: Software for the Statistical Analysis of Correlated Data, release 9.1. Triangle Park, NC: Research Triangle Institute.

Thomsen, I. B. and Siring, E. (1983), "On the Causes and Effects of Nonresponse: Norwegian Experiences," in W. G. Madow and I. Olkin (eds.), Incomplete Data in Sample Surveys, Vol. 3, New York: Academic Press.

Triplett, T., Blair, J., Hamilton, T. et al. (1996), "Initial Cooperators vs. Converted Refusers: Are there Response Behavior Differences?" In: 1996 Proceedings of the Section on Survey Research Methods, American Statistical Association. Alexandria, VA: American Statistical Association, 1996:1038-41.

Verba, S. (1996), "The Citizen as Respondent:Sample Surveys and American Democracy," American political Science Review, 90:1-7.

Voigt, L. F., Koepsell, T. D., and Darling, J. R. (2003), "Characteristics of Telephone Survey Respondents According to Willingness to Participate," American Journal of Epidemiology, 157:66-73.

Yan, T., Curtin, R., Jans, M. (2010), "Trends in Income Nonresponse Over Two Decades," Journal of Official Statistics, 26(1):145-164.