

Nonresponse Bias Analysis in IRS' Survey of Low-income Taxpayersⁱ

Karen Masken¹, Kerry Levin², Cynthia Helba², and Katie Hubbell²
 Internal Revenue Service¹
 Westat²

1. Introduction

While almost all types of surveys have suffered from increasing non-response in the last decade (Groves, Fowler, Couper, Lepkowski, Singer, and Tourangeau, 2004), surveys of low-income households present some particular response issues. For example, some households do not have telephones, and of those who do, many do not have land lines. If those without telephones or with only cellular phones have different response patterns than households with land lines, then typical telephone-only survey estimates will naturally be biased. Low income households also appear to be a relatively mobile group and therefore harder to locate if a list sample is being used (Judkins, Shapiro, Brick, Flores-Cervantes, Ferraro, Strickler, and Waksbery, 1999). Again, if those who cannot be located differ in their responses from those who were located, bias will be introduced. A survey of low-income taxpayers was recently sponsored by the Internal Revenue Service (IRS) and conducted by Westat. This paper describes the survey methods used to maximize response rates among this unique population. Next, the analysis done to evaluate any non-response bias is described, followed by a discussion of potential adjustments that could be made based on the analysis. Finally, recommendations for future surveys of this nature are discussed.

2. Background

The Earned Income Tax Credit (EITC) was created in 1975 in an effort to provide financial support to working families and individuals with low incomes. Eligibility and the amount of the credit received are largely determined by income level and whether or not the taxpayer claims "qualifying children." Children claimed for the EITC are considered to be qualifying if they meet certain tests based on age, relationship, and residency. Although the EITC helps millions of low-income working Americans, it has also been subject to misunderstanding and abuse on the part of some taxpayers. An IRS study of 1999 EITC claims revealed that overpayments comprised more than one-quarter of all claims, amounting to \$8.5 to \$9 billion in un-recovered funds. For those who made qualifying child errors, the most common error was claiming

children for the EITC who failed to meet the residency test.

To help reduce erroneous claims, IRS began testing a program with taxpayers that requires them to provide certification that children claimed under the EITC actually meet the residency test before their EITC claims are accepted and refunds paid. The sample for this test consisted of a test group and a control group. Each included 25,000 taxpayers who were from groups that the IRS could not systemically determine had a low likelihood of claiming children who do not meet the residency requirements. All respondents in the test group were sent a letter in December 2003 informing them that in order to claim the EITC with qualifying children for tax year 2003, they would be required to complete a Qualifying Children Residency Statement and provide documentation as part of their income tax return. The control group had similar characteristics to those in the test group, but was not asked to provide additional documentation regarding residency. As part of the evaluation of this test, IRS sponsored a survey conducted by Westat of these low-income taxpayers in order to understand their experiences with, and attitudes about, the certification program.

3. Survey Design and Administration

Separate surveys and survey samples were designed for the test and control groups. The sample size for the test group was 11,826 and the sample size for the control group was 2,068. A multi-mode data collection was employed in each group. The primary mode of data collection was a computer-assisted telephone interviewing (CATI). For those taxpayers who did not have a listed telephone number, paper-and-pencil mail surveys were delivered via first class mail. The mail survey of taxpayers for whom telephone numbers could not be located was conducted to reduce the potential for bias related to the availability of telephone numbers. (Analysis performed at the end of the survey established that the mail respondents did in fact have had different characteristics than the phone respondents (Levin, Helba, Forsyth and Masken, 2005).) Prenotification letters, which included a token incentive (magnet), were sent to all sampled taxpayers. For the telephone survey, an automated call scheduling

system divided the week into time slices and moved each case through these time periods in a specified pattern of call attempts. During the data collection period, additional mail surveys were sent to cases where telephone contact was made with a household member and that member stated they did not know the sampled taxpayer or cases where a telephone company recording stated that the number dialed was no longer working or was temporarily disconnected. In addition, refusal conversion letters were sent to interim refusal cases, maximum call cases and other interim cases (e.g., ring-no-answer; busy; answering machine).

4. Response Rates

Table 1 presents the response and cooperation rates. The overall response rate was quite low, 24 percent. It is interesting to note that the response rates for those in the test group who claimed EITC with qualifying children were significantly higher than the other groups of respondents. This is likely due to the fact that the topic was salient to this group of taxpayers (they were the only group subject to the certification test; those in the test group who did not claim EITC with qualifying children were excluded). Literature suggests that respondents for whom the topic is salient are more likely to participate (Groves and Couper, 1998). The other note of interest is that the cooperation rates were quite high. In other words, once contact was made, many respondents were willing to participate. This suggests that the low response rate was due primarily to not being able to locate respondents, not to refusals.

There were three types of respondents who could not be located during data collection: (1) cases where no contact was ever made (maximum of 25 calls made), (2) non-working telephone numbers or returned mail, and (3) cases where telephone contact was made with a household member and that member stated they did not know the sampled taxpayer. However, given the unusually low refusal rate, it is highly likely that the last group of non-locatables included a number of ‘soft’ refusals. That is, rather than tell the interviewer they did not wish to cooperate, the respondent simply stated the sampled tax payer did not live there. Soft refusals such as this situation would make the cooperation rate artificially high.

5. Nonresponse Analysis

To assess the degree of bias presented by such high levels of nonresponse, characteristics of respondents and nonrespondents were compared using IRS administrative data. The comparisons between respondents and nonrespondents were tested for statistical significance using Student’s T-tests. The difference was deemed significant if the p-value was less than 0.05.

Table 2 presents these comparisons for the test and control groups. For both the test and control groups, taxpayers who did not file a tax return responded at a significantly lower rate than those who did file a tax return. For those who did not file a return during the study year, the address information IRS has would be from the prior year. The age of the address data is probably the main reason those who did not file a tax return in the study year were difficult to locate.

Table 1. Response and Cooperation Rates for Sampled Respondents

	Survey Sample Size	Response Rate	Cooperation Rate
Overall	13,171	24%	78%
Test Group			
Claimed EITC with Qualifying Children	2,729	33%	85%
Did not Claim EITC with Qualifying Children	8,374	21%	77%
Control Group			
Claimed EITC with Qualifying Children	478	22%	77%
Did not Claim EITC with Qualifying Children	1,590	20%	72%

For those who filed a return in the test group, respondents who claimed EITC with qualifying children responded at a significantly higher rate than those who did not claim the EITC with qualifying children. The same does not hold true for the control group where the EITC claimants and non-claimants responded at the same rate. As noted earlier, the likely reason for this is topic saliency. Those who had to go through the certification process had a vested interest in the topic and wanted to talk about their experience.

For both the test and control groups, there are significant differences in response rates among age groups. As is typical with household surveys, younger taxpayers are less likely to respond than older ones (Groves and Couper, 1998). This is probably due to younger people being out and away from home more often than older people and therefore more difficult to contact. Also, younger taxpayers are probably more mobile and likely to rely solely on cellular telephones than older ones making them more difficult to locate. Conversely, a comparison between regions yields no differences in response rates.

Table 2. Analysis of Respondents and Nonrespondents

	Test Group			Control Group		
	Survey Sample	Respondents	Non-respondents	Survey Sample	Respondents	Non-respondents
Did not File a Return	31%	19%	35%*	35%	26%	37%*
Filed a Return	69%	81%	65%*	65%	74%	63%*
<i>Of Filers</i>						
Claimed EITC With Qualifying Children	25%	33%	22%*	23%	25%	23%
No EITC Claim with Qualifying Children	75%	67%	78%*	77%	75%	77%
<i>Age</i>						
Under 31 years old	30%	22%	33%*	30%	20%	33%*
31 under 41	27%	24%	28%*	28%	29%	27%*
41 under 51	26%	29%	24%*	26%	29%	25%*
51 years and older	18%	24%	15%*	17%	22%	15%*
<i>Region</i>						
MW	17%	19%	17%	18%	20%	17%
NE	14%	14%	14%	15%	16%	14%
SO	52%	52%	53%	49%	46%	50%
WE	16%	16%	16%	18%	18%	18%
Married Filing Jointly	13%	16%	12%*	15%	22%	13%*
Not Married Filing Jointly	87%	84%	88%*	85%	78%	87%*
<i>Of non-married Filers</i>						
Male	67%	31%	67%*	71%	64%	70%*
Female	33%	39%	33%*	29%	36%	30%*

* P<0.05

The next comparison looks at the difference in response rates between taxpayers who used the filing status of “married filing jointly” versus those who did not (generally single or head of household filers). For both the test and control groups, the “married filing jointly” taxpayers responded at a significantly higher rate than those who did not file joint tax returns. This is typical of household surveys. The more adults there are in a household, the higher the probability that that someone will be home to respond to the survey.

Of the non-married respondents, there is a significant difference in both the test and control groups between male and female respondents. Females responded at a significantly higher rate. This is not an unexpected finding as women tend to respond to surveys at a higher rate than men (Smith, 1983).

6. Nonresponse Adjustment

As illustrated above, non-respondents had different characteristics than respondents. If responses to survey questions are correlated with these characteristics, then the survey results will be biased unless adjustments are made to account for non-response bias. For this survey, it was determined that the weights should be adjusted to reduce the bias. Chi-squared Automatic Interaction Detector (CHAID) analysis was used to determine the best adjustment cells. CHAID analysis divides the population into two or more distinct groups based on categories of the “best” predictor of a dependent variable. In this case the dependent variable was the response rate.

Table 3. Comparison of Survey Characteristics after adjustments to the Study Population

	Test Group		Control Group	
	Westat Survey	IRS Population	Westat Survey	IRS Population
Did not File a Return	14%	11%	12%	10%
Filed a Return	86%	89%	88%	90%
<i>Of Filers</i>				
Claimed EITC With Qualifying Children	66%	67%	73%	74%
No EITC Claim with Qualifying Children	34%	33%	27%	26%
<i>Age</i>				
Under 31 years old	30%	30%	29%	29%
31 under 41	29%	29%	28%	28%
41 under 51	25%	25%	26%	26%
51 years and older	16%	16%	17%	16%
<i>Region</i>				
MW	16%	16%	17%	16%
NE	14%	14%	16%	14%
SO	53%	52%	51%	52%
WE	17%	18%	17%	18%
Married Filing Jointly	11%	11%	13%	10%
Not Married Filing Jointly	89%	89%	87%	90%
<i>Of non-married Filers</i>				
Male	64%	64%	63%	65%
Female	36%	36%	37%	35%

The CHAID analysis determined that the best predictor of response rate was whether a phone number was found for the respondent or not. The population was split into these two groups (i.e., phone number located) and then split further into subgroups based on age. This splitting continued until no more statistically significant differences in response rate by age could be found. The weights for responding persons within each of the newly formed nonresponse cells were then adjusted to compensate for nonresponding persons. To the extent that survey questions are answered similarly by people with a nonresponse cell, the application of such nonresponse adjustment factors reduces the bias due to nonresponse. After the non-response adjustment, poststratification was performed as well using filing status, gender, region and EITC claims as the poststratification variables.

Table 3 presents comparisons of characteristics between the weighted survey data after adjustments and the IRS population based on administrative data. Since most of the variables presented here were used in either non-response adjustment or poststratification, it is not surprising that the survey distributions mirror the population distributions (because the survey sample for the control group was relatively small, it doesn't match the population exactly). What are of interest are the two variables that were not used in adjustment. Whether the respondent filed a return or not was not an adjustment variable, but the bias has been reduced considerably after the adjustments. Likewise, whether or not the taxpayer claimed EITC was not used in adjustment but also has a distribution more similar to the population than it did prior to the adjustments.

7. Conclusion

Steps described in the Survey Design and Administration section were valuable in increasing response rates and thereby reducing the potential for nonresponse bias. The weighting adjustments used further reduced the bias for the variables used in adjustment and significantly lowered it for other key variables. However, IRS has very limited demographic information regarding taxpayers. This lack of information necessarily limits the extent of nonresponse analysis possible. Therefore, the survey may have unobservable bias related to nonresponse. For example, the IRS has no information on the race/ethnicity of taxpayers – a variable that is customarily used in non-response adjustment. If race/ethnicity is a predictor of response for this particular survey, then there may be some response bias that cannot be observed. Future taxpayer studies should include nonresponse bias components that

focus on subsamples of nonrespondents. Such studies would allow researchers to capture information that would further elucidate biases related to non response in taxpayer surveys.

References

- Groves, R., and Couper, M., *Nonresponse in Household Interview Surveys*, John Wiley & Sons, Inc. New York, NY, 1998.
- Groves, R., Dillman, D., Eltinge, J., and Little, R., *Survey Nonresponse*, John Wiley & Sons, Inc. New York, NY, 2002.
- Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R., *Survey Methodology*, John Wiley & Sons, Inc. New Jersey, 2004.
- Judkins, D., Shapiro, G., Brick, M., Flores-Cervantes, I., Ferraro, D., Strickler, T., and Waksberg, J. (1999). *NSAF Methodology Reports, No 2, NSAF Sample Design*, Washington, DC: Urban Institute, 1997.
- Levin, K., Helba, C., Forsyth, B., and Masken K., What's the benefit of a mail supplement to a telephone survey for gathering data from "hard to reach" populations? Presented at the 60th Annual Meeting of the American Association for Public Opinion Research, Miami Beach, FL, May 2005.
- Smith, T., "The hidden 25 percent: An analysis of nonresponse on the 1980 General Social Survey", *Public Opinion Quarterly*, 47, pp 386-404, 1983.

ⁱ The views expressed in this paper are the authors and do not necessarily represent the position of the Internal Revenue Service.