

2009 National Health Interview Survey Sample Adult and Sample Child nonresponse bias analysis

by Chris Moriarity, Division of Health Interview Statistics, National Center for Health Statistics

Introduction

Nonresponse occurs in a survey when information is not collected about an eligible sample member. There are several types of nonresponse in the National Health Interview Survey (NHIS). Unit nonresponse, or household-level nonresponse, occurs when no information is collected about any members of a sample NHIS household. Item nonresponse occurs when information is not collected for a questionnaire item. A third type of nonresponse occurs in the NHIS when information is not collected for an entire section, e.g., the Sample Adult Section, or the Sample Child Section.

The annual NHIS Survey Description Document (see References), Appendix I, Table I, provides unit and section response rates. The 2009 unit response rate is 82.2%. The 2009 conditional section response rates vary from 99.3% (family) to 89.9% (sample child) to 80.1% (sample adult).

Nonresponse can lead to bias in survey estimates if the characteristics of the nonrespondents are different from the characteristics of the respondents, and if weight adjustments for nonresponse do not adequately account for the difference.

The household weights contain a geographic nonresponse adjustment to compensate for the household-level nonresponse. The household weight nonresponse adjustment assumes that the characteristics of responding and nonresponding households that are geographically close to each other are similar.

The focus of this nonresponse bias analysis is on the nonresponse to the 2009 Sample Adult and Sample Child Sections. The findings from this analysis led to the decision to add a geographic nonresponse adjustment to the Sample Adult and Sample Child File weights, beginning with the 2010 data. The Sample Adult File analysis is presented first, followed by the Sample Child File analysis.

NHIS 2009 Sample Adult File Analysis

The final response rate for the 2009 Sample Adult File is 65.4%, and the conditional response rate for the 2009 Sample Adult File is 80.1%. The final response rate takes account of nonresponse at the household and family levels. The conditional response rate for the Sample Adult File is the response rate for the Sample Adult Section, given that household and family response occurred. The final response rate of 65.4% for the 2009 Sample Adult File is obtained by multiplying 82.2% (the household response rate), 99.3% (the family response rate), and 80.1% (the conditional sample adult rate).

All nonrespondents to the Sample Adult Section are present in the Person File, i.e., the Person File contains person-level data for the sample adult nonrespondents. The respondents to the Sample Adult Section also are present in the Person File. The Person File also contains a variable indicating whether a person was selected to be the sample adult, and if so, whether the Sample Adult Section was completed. Thus, it is possible to do a comparison of Person File

characteristics for sample adult respondents and sample adult nonrespondents.

The Sample Adult Section response rate differs by whether or not the family respondent is selected as the sample adult (Dahlhamer and Simile, 2009 - see References). In 2009, 88% of family respondents selected as sample adults completed the Sample Adult Section, but only 65% of non-family respondents selected as sample adults completed the Sample Adult Section. Persons selected as sample adults who were not the family respondent are more likely to be male, younger, Hispanic, married, and less likely to be poor. However, persons selected as sample adults are similar in current health status, regardless of being the family respondent or not.

Table 1 presents a comparison of sample adult respondents and sample adult nonrespondents on selected Person File variables. The comparisons are shown separately by whether the family respondent was selected as the sample adult or not. The estimates are weighted estimates, using the Person File weight that does not include the final adjustment to independent population estimates that is described in more detail below.

Table 1: Percent Distributions, using Person File data, of Selected Demographic and Health Characteristics for Sample Adult Section Respondents and Nonrespondents

	Family respondent			
	YES		NO	
	Responded to Sample Adult Section			
	YES	NO	YES	NO
Sex:				
Male	41%	45%	54%	56%
Female	59%	55%	46%	44%
Age:				
18-44	42%	49%	48%	57%
45-64	36%	31%	34%	34%
65+	22%	20%	18%	9%
Hispanic origin:				
Hispanic	10%	12%	14%	15%
Non-Hispanic	90%	88%	86%	85%
Poverty status (1):				
Poor	15%	11%	8%	7%
Near poor	17%	10%	14%	12%
Not poor	59%	38%	71%	61%
Unknown	9%	41%	7%	20%
Metropolitan area size (2):				
Large MSA	52%	60%	49%	58%
Small MSA	32%	27%	33%	29%
Non-MSA	16%	13%	18%	13%
Census Region (3):				
Northeast	18%	18%	17%	20%
Midwest	25%	18%	25%	20%
South	36%	42%	36%	36%
West	21%	22%	22%	24%
Marital status:				
Married	36%	38%	70%	64%
Other	64%	62%	30%	36%
Current health status:				
Excellent/Very Good	59%	59%	60%	65%
Good	27%	27%	27%	26%
Fair/Poor	14%	13%	13%	9%
Unknown	-	1%	-	-

-: less than 0.5%

(1): Poverty status is based on family income and family size using the U.S. Census Bureau's poverty thresholds for the previous calendar year. "Poor" persons are defined as below the poverty threshold. "Near poor" persons have incomes of 100% to less than 200% of the poverty threshold. "Not poor" persons have incomes that are 200% of the poverty threshold or greater. Because of the different income questions used in 2007 and beyond, poverty ratio estimates may not be comparable with those from earlier years.

(2): MSA is metropolitan statistical area (June 6, 2003 definitions). Large MSAs have a population size of 1 million or more; small MSAs have a population size of less than 1 million. "Non-MSA" consists of persons not living in a metropolitan statistical area.

(3): Census Region: in the geographic classification of the U.S. population, states are grouped into four regions by the U.S. Census Bureau. Northeast: Maine, Vermont, New Hampshire, Massachusetts, Connecticut, Rhode Island, New York, New Jersey, and Pennsylvania. Midwest: Ohio, Illinois, Indiana, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Kansas, and Nebraska. South: Delaware, Maryland, District of Columbia, West Virginia, Virginia, Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, Louisiana, Oklahoma, Arkansas, and Texas. West: Washington, Oregon, California, Nevada, New Mexico, Arizona, Idaho, Utah, Colorado, Montana, Wyoming, Alaska, and Hawaii.

Informally (no hypothesis testing was done), Table 1 indicates that nonrespondents are more likely to be younger and Hispanic. Respondents and nonrespondents have similar current health status.

The final step in the creation of the 2009 sample adult weight and the sample child weight is a ratio adjustment to independent population estimates created by the U.S. Bureau of the Census. The ratio adjustments are applied by sex, age groups, and race/ethnicity groups. Thus, this ratio adjustment helps to alleviate potential nonresponse bias by sex, age, and race/ethnicity.

The breakout by response/nonresponse and whether the family respondent is selected as the sample adult or not shows some geographic differences. Nonresponse is higher in large MSAs, and response is higher in the Midwest Census Region.

The sample adult weight for 1997-2009 does not contain a geographic nonresponse adjustment factor. When the sample adult weighting specifications were developed for the first time for the 1997 NHIS, the level of nonresponse to the sample adult section was unknown. The assumption that was made then was that nonresponse would be taken care of adequately by the ratio adjustment to the independent population estimates.

Table 2 shows geographic population proportion estimates from three sources. The first column shows the estimates from the Sample Adult File, using the final Sample Adult File weight. The third column shows the estimates for all persons from the Person File, using the final Person File weight. The middle column shows estimates from the Sample Adult File, using an experimental final weight that includes a geographic nonresponse adjustment, computed in the same way as the household weight geographic nonresponse adjustment (see pp. 15-17 of Series 2, No. 130 - see References).

Table 2: Percent Distributions, using Sample Adult File and Person File data, of Metropolitan Area Size and Census Region

	Using Sample Adult File final weight	Using Sample Adult File experimental final weight	Using Person File final weight
Metropolitan area size:			
Large MSA	51.6%	52.7%	53.2%
Small MSA	31.9%	31.4%	31.0%
Non-MSA	16.5%	15.9%	15.8%
Census Region:			
Northeast	17.5%	17.9%	18.0%
Midwest	24.2%	23.5%	23.1%
South	35.8%	35.9%	36.2%
West	22.5%	22.7%	22.7%

The estimates based on the Person File are considered to be the benchmark, because they are based on a larger number of sample persons. It is apparent that the geographic nonresponse adjustment makes the sample adult estimates more similar to the Person File estimates. For all seven estimates, the geographic nonresponse adjustment moves the Sample Adult File estimate towards the Person File estimate. It compensates for the higher nonresponse in large MSAs, and the lower nonresponse in the Midwest Census Region (see Table 1).

The comparisons shown in Table 2 were also carried out for 2006, 2007, and 2008. The results are similar to what is shown in Table 2.

We examined the effect of the experimental final weight on selected estimates that appear

in Appendix III of Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2009 (see References). No meaningful differences were found. Consistent with what is shown in Table 2, large MSA estimates tended to go up slightly, and small MSA estimates tended to go down slightly. Also, Northeast estimates tended to go up slightly, and Midwest estimates tended to go down slightly.

NHIS 2009 Sample Child File Analysis

The final response rate for the 2009 Sample Child File is 73.4%, and the conditional response rate for the 2009 Sample Child File is 89.9%. The final response rate takes account of nonresponse at the household and family levels. The conditional response rate for the Sample Child File is the response rate for the Sample Child Section, given that household and family response occurred. The final response rate of 65.4% for the 2009 Sample Child File is obtained by multiplying 82.2% (the household response rate), 99.3% (the family response rate), and 89.9% (the conditional sample child rate).

For convenience, we refer below to "Sample Child Section respondents", "Sample Child Section nonrespondents", etc. It is not literally true that the sample children are respondents or nonrespondents, because the sample child data are provided by an adult.

All nonrespondents to the Sample Child Section are present in the Person File, i.e., the Person File contains person-level data for the sample child nonrespondents. The respondents to the Sample Child Section also are present in the Person File. The Person File also contains a variable indicating whether a person was selected to be the sample child, and if so, whether the Sample Child Section was completed. Thus, it is possible to do a comparison of Person File characteristics for sample child respondents and sample child nonrespondents.

Table 3 presents a comparison of sample child respondents and sample child nonrespondents on selected Person File variables. The estimates are weighted estimates, using the Person File weight that does not include the final adjustment to independent population estimates that was described in more detail above.

Table 3: Percent Distributions, using Person File data, of Selected Demographic and Health Characteristics for Sample Child Section Respondents and Nonrespondents

	Respondent	
	YES	NO
Sex:		
Male	50.5%	52.4%
Female	49.5%	47.6%
Age:		
18-44	27.8%	25.7%
45-64	35.6%	34.9%
65+	36.6%	39.4%
Hispanic origin:		
Hispanic	20.9%	22.4%
Non-Hispanic	79.1%	77.6%
Poverty status:		
Poor	16.1%	8.4%
Near poor	19.1%	13.4%
Not poor	58.1%	40.3%
Unknown	6.7%	37.9%
Metropolitan area size:		
Large MSA	53.1%	58.6%
Small MSA	31.5%	27.9%
Non-MSA	15.4%	13.5%
Census Region:		
Northeast	17.3%	17.6%
Midwest	23.5%	15.8%
South	35.9%	39.8%
West	23.3%	26.8%
Current health status:		
Excellent/Very Good	84.4%	82.4%
Good	13.9%	16.0%
Fair/Poor	1.7%	1.4%
Unknown	-	0.2%

-: less than 0.5%

Informally (no hypothesis testing was done), Table 3 indicates that respondents and nonrespondents have minor differences by sex, age, and Hispanic origin. Respondents appear to be a little healthier than nonrespondents. Nonresponse is higher in large MSAs, and response is higher in the Midwest Census Region.

The sample child weight for 1997-2009 does not contain a geographic nonresponse adjustment factor. When the sample child weighting specifications were developed for the first time for the 1997 NHIS, the level of nonresponse to the sample child section was unknown. The assumption that was made then was that nonresponse would be taken care of adequately by the ratio adjustment to the independent population estimates.

Table 4 shows geographic population proportion estimates from three sources. The first column shows the estimates from the Sample Child File, using the final Sample Child File weight. The third column shows the estimates for all persons from the Person File, using the final Person File weight. The middle column shows estimates from the Sample Child File, using an experimental final weight that includes a geographic nonresponse adjustment, computed in the same way as the household weight geographic nonresponse adjustment.

Table 4: Percent Distributions, using Sample Child File and Person File data, of Metropolitan Area Size and Census Region

	Using Sample Child File final weight	Using Sample Child File experimental final weight	Using Person File final weight
Metropolitan area size:			
Large MSA	52.5%	52.8%	52.9%
Small MSA	31.6%	31.4%	31.2%
Non-MSA	15.9%	15.8%	15.9%
Census Region:			
Northeast	16.9%	17.0%	16.8%
Midwest	24.1%	23.8%	23.4%
South	35.1%	35.0%	35.6%
West	23.9%	24.2%	24.2%

The estimates based on the Person File are considered to be the benchmark, because they are based on a larger number of sample persons. It is apparent that the geographic nonresponse adjustment makes the Sample Child File estimates more similar to the Person File estimates. For four out of seven estimates, the geographic nonresponse adjustment moves the Sample Child File estimate towards the Person File estimate. It compensates for the higher nonresponse in large MSAs, and the lower nonresponse in the Midwest Census Region (see Table 3).

The comparisons shown in Table 4 were also carried out for 2008. The results are similar to what is shown in Table 4. For five out of seven estimates, the geographic nonresponse adjustment moves the Sample Child File estimate towards the Person File estimate.

We examined the effect of the experimental final weight on selected estimates that appear in Appendix III of Summary Health Statistics for U.S. Children: National Health Interview Survey, 2009 (see References). The experimental weight usually gave slightly different estimates, but no meaningful differences were found, and there was no clear pattern in the changes.

Summary

It is possible to make some detailed comparisons of 2009 Sample Adult Section respondents and Sample Adult Section nonrespondents using 2009 Person File data. Differences in response rates by sex, age, and race/ethnicity are being compensated for by the ratio adjustment to independent population estimates.

Differences in geographic response rates have been identified by this analysis, and a methodology is being implemented in the 2010 NHIS to compensate. The Division of Health Interview Statistics is going to include a geographic nonresponse adjustment to the sample adult weight and the sample child weight, beginning with the 2010 NHIS. This will alleviate nonresponse bias due to differences in response rates by geographic area.

References

Bloom B, Cohen RA, Freeman G. Summary Health Statistics for U.S. Children: National Health Interview Survey, 2009. National Center for Health Statistics. Vital Health Stat 10(247). 2010.

Botman SL, Moore TF, Moriarity CL, Parsons VL. Design and Estimation for the National Health Interview Survey, 1995-2004. National Center for Health Statistics. Vital Health Stat

2(130). 2000.

Dahlhamer JM, Simile CM. Subunit Nonresponse in the National Health Interview Survey (NHIS): An Exploration Using Paradata. *ASA Proceedings of the Joint Statistical Meetings*, 262-276, 2009.

National Center for Health Statistics. 2009 National Health Interview Survey (NHIS) Public Use Data Release: NHIS Survey Description. National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Maryland. 2010.

Pleis JR, Ward BW, Lucas JW. Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2009. National Center for Health Statistics. *Vital Health Stat* 10(249). 2010.