UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

Final DSMD Signed

SEP 2 2014

MEMORANDUM FOR Jason Fields
Survey Director, Survey of Income and Program Participation
Associate Directorate for Demographic Programs

From: Ruth Ann Killion
Chief, Demographic Statistical Methods Division

Subject: Nonresponse Bias Analysis for Wave 2 2008 Survey of Income and Program Participation (SIPP) (ALYS-13)

This document describes the results of the nonresponse bias analysis for Wave 2 of the 2008 Survey of Income and Program Participation (SIPP).

If you have questions, please contact Tracy Mattingly at 301-763-6445 or via email at Tracy.L.Mattingly@census.gov or Jamie Choi at 301-763-4554 or via email at Jamie.Choi@census.gov.

Attachment

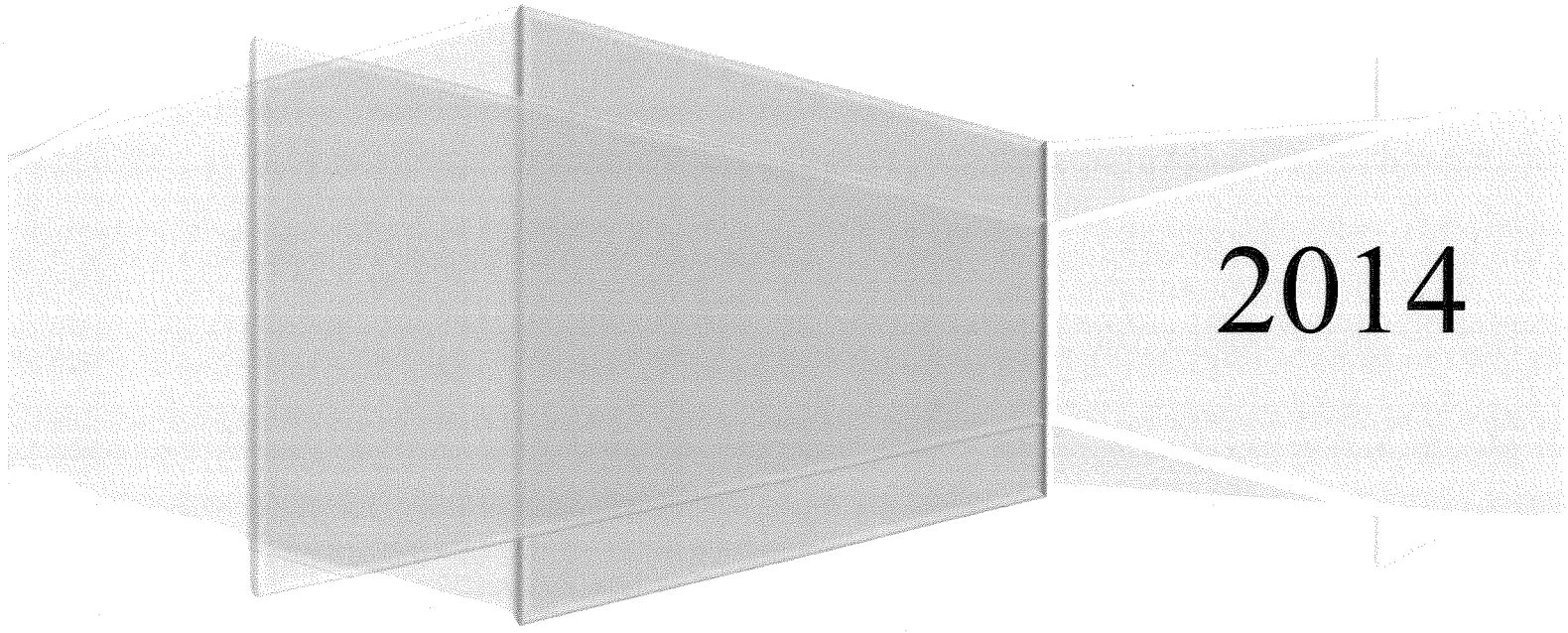
cc:

D. Doyle (ADDP)
M. Marlay
V. Velkoff (SEHSD)
G. Gathright
J. Farber (DSMD)
J. Hill
J. Scott
T. Mattingly
J. Choi
R. Culver III
S. Mack
M. Sundukchi
A. Westra
E. Wong

U.S. Census Bureau

Nonresponse Bias Analysis for Wave 2 of the 2008 Survey of Income and Program Participation (SIPP)

By Jamie Choi and Tracy Mattingly



2014

Introduction

For a household survey, unit nonresponse is the failure to obtain any survey measures on a sample unit (e.g. household). These nonresponse rates have been increasing in recent years, even for the large government surveys. With this increasing nonresponse there is also growing concern over data quality and losing valuable information from these nonrespondents. Declining response rates can be an indicator of nonresponse bias, or a difference in survey measures between respondents and nonrespondents, which can affect data quality. However, there is not always a direct link between response rates and nonresponse bias since nonresponse bias can vary across different statistics in the same survey; low response rates in surveys may yield to some statistics having large nonresponse bias [3]. Therefore, understanding and measuring nonresponse bias for key estimates is an important aspect in determining overall data quality.

Because policy makers use estimates from the demographic surveys conducted by the U.S. Census Bureau and other agencies to determine the successfulness of programs or for national economic indicators, the highest data quality is necessary. This has led more surveys to investigate nonresponse bias, especially after the Office of Management and Budget (OMB) released their standards in 2006 that require survey programs to make plans for a nonresponse bias analysis if unit response rates fall below 80% [7]. The Census Bureau incorporated this guideline into its own standards along with guidelines which state that serious data quality issues related to nonsampling error can occur when cumulative response rates for a longitudinal survey fall below 60% and when sample attrition from wave to wave is greater than 5% [15].

The 2008 Survey of Income and Program Participation (SIPP) obtained a response rate of 80.8% at the first interview and 74.2% at the second interview. The panel ended in December 2013 with a cumulative response rate of 49.4%.

This report presents an analysis of nonresponse bias for Wave 2 of the 2008 SIPP. The analysis takes into account the longitudinal aspect of the survey and utilizes the data available for those who drop out after Wave 1 to give a better picture of the overall bias associated with nonresponse. Three different studies are used to examine potential nonresponse bias. The methods include comparing key estimates of the full sample to the respondent sample, analyzing estimates by predicted response propensity quintiles, and calculating the representivity of the survey using R-indicators.

Previous Research

Previous efforts to examine the bias and determine the impact on SIPP estimates included comparing characteristics of households that responded in all waves versus those that dropped out. The SIPP nonresponse workgroup found that households that were renting, living in large urban areas, and had young adults (15-24) as the householder were more likely to be nonrespondents [12]. Another study compared SIPP annual poverty rates and health insurance coverage to CPS-ASEC. Results showed differences in poverty rates at the 150% and 200% poverty thresholds and differences in health insurance coverage, especially among blacks and

Hispanics [9]. Mack and Petroni [4] summarized the results from several studies that looked into using logistic regression and various raking methodologies for nonresponse weighting adjustments. The results indicated that alternative weighting procedures did not reduce the bias more than the current adjustment for estimates of income, unemployment, government assistance and poverty.

A study completed on Wave 1 of the 2008 Panel showed through benchmark analysis that the SIPP is underestimating participation in Supplemental Nutrition and Assistance Program (SNAP), Supplemental Security Income (SSI), Medicaid, and Medicare compared to the program sources. Examining response rates across different subgroups found that those living in the Northeast had the lowest response rate out of all regions, Black householders had a lower response rate compared to Nonblack householders, and metropolitan statistical area (MSA) and principal city residents had lower response rates compared to those not living in the MSA or in the principal city. Comparing estimates of the frame variables for the full sample and for respondents-only revealed differences for region, urban/rural status, CBSA type, and race. The study also showed through odds ratios obtained from a logistic regression analysis that region, household size, and age of householder significantly affected the likelihood of responding [5].

Data

The SIPP is a longitudinal survey that collects detailed information about income, employment, health insurance, and program participation for the civilian, noninstitutionalized population living in the United States. The 2008 SIPP panel started on September 2008 and ended on December 2013 with interviews being conducted at four-month intervals called waves.

The Census Bureau employed a two-stage sample design to select the 2008 SIPP sample. A systematic selection was used to select housing units within 351 primary sampling units (PSUs) from the master address file created from the 2000 Decennial Census. In addition, households located in areas with a higher concentration of low-income households were oversampled by 44 percent to increase the accuracy of estimates for statistics of low-income households and program participation [13].

When a respondent is interviewed, data is collected about the four preceding months. These four reference months comprise one wave. The sample in Wave 1 of the 2008 SIPP consisted of approximately 65,500 households of which only 52,030 of the households were eligible for interview. Of those eligible households, 42,030 were interviewed, with a response rate of 80.8% [17]. In subsequent waves, all adults who were interviewed in Wave 1 were followed and interviews were attempted for all household members.

Because SIPP is longitudinal, single wave response rates do not give a complete picture of nonresponse. The total sample attrition is also examined. The SIPP measures sample attrition using a sample loss rate, which is calculated by taking into account the Wave 1 nonresponse and each subsequent wave's nonresponse [17]. Given that SIPP interviews new household members joining the household after Wave 1, the growth of nonresponding households is estimated and included in the sample loss rate. The sample loss rate for Wave 1 is 19.2%. For Wave 2, the sample loss rate is 25.8% and is 49.4% for the final wave of the panel.

This study analyzes nonresponse bias for Wave 2 of the 2008 SIPP. The sample (i.e., households that were interviewed in Wave 1) for Wave 2 is 42,029[1] households. Of those households, 37,471[2] were interviewed in Wave 2.

Analytic Variables

This report focuses on analysis of nonresponse bias for key estimates of the 2008 SIPP. Actual estimates of nonresponse bias can only be produced for variables that are known for both respondents and nonrespondents. Because of the longitudinal aspect of the survey, information on key estimates collected in Wave 1 can be used to analyze nonresponse bias in later waves of the survey.

The SIPP key estimates that are assessed for potential nonresponse bias include:

- *Mean (monthly) household earnings*
- *Mean (monthly) household income*
- *Percent of households where at least one household member was covered by Medicaid*
- *Percent of households where at least one household member was covered by Medicare*
- *Percent of households where at least one household member received Social Security*
- *Percent of households where at least one household member received SSI*
- *Percent of households where at least one household member received SNAP benefits*

In addition, variables used for household noninterview weighting adjustments are also examined [6]. These variables include:

- *Age of Reference Person* (Four levels: Under 25, 25-34, 35-54, 55+)
- *Assets* (Two levels: Bonds/Etc. - at least one HH member possessed at least one of the following assets: money market deposit accounts, certificates of deposit, mutual fund shares, rental property, mortgages, royalties, or other financial investments; Minimal – other)
- *CBSA area* (Three levels: In Principal City of MSA, In MSA not in Principal City, Not in MSA or Principal city)
- *Education* - Highest level of school completed or the highest degree received by the reference person (Four levels: Less than high school; 9th-12th, no diploma; High school graduate, some college (no degree), vocational/technical school, associate degree; Bachelor's/Master's/Doctorate/Professional degree)
- *Gender of Reference Person* (Two levels: Male, Female)
- *Household Size* (Four levels: 1 persons in the household, 2 persons in the household, 3 persons in the household, 4+ persons in the household)

[1] The analysis does not include households with reference persons whose age is less than 15. The number of interviewed households including those households is 42,030.

[2] The analysis in this report does not include households spawned from an original sample household. The number of interviewed households including spawned households for Wave 2 is 38,998.

- *Household Type* (Three levels: Reference person is female with no husband present and with her own children less than 16 years old (FHHNSP), Reference person is 65 years old or older, Other)
- *Income Type* (Two levels: Welf/Etc. - at least one household member received income from at least one of the following sources: SSI, Aid to Families with Dependent Children (AFDC), other welfare, Women, Infants, and Children Nutrition Program (WIC), SNAP, or Medicaid; Other)
- *Race of Reference Person* (Two levels: Black, Non-Black)
- *Region* (Four levels: Northeast, Midwest, South, West)
- *Tenure* (Two levels: Renter, Owner)
- *Urban/Rural* (Two levels: Urban area, Rural area)
- *Within PSU Strata* (Two levels: Low income strata, Non-low income strata) This variable was based on a probability calculated using 2000 Census Long Form income data within a designated geographic region.

Methods

The analysis involves using three different studies for measuring nonresponse bias for key estimates in Wave 2 of the 2008 SIPP. These studies utilize Wave 1 data of the 2008 Panel to examine possible nonresponse bias occurring in Wave 2 of the SIPP. The methods include comparing estimates of the full sample to the respondent sample, analyzing estimates by predicted response propensity quintiles, and calculating the representivity of the survey using R-indicators.

Comparing Estimates of the Full Sample to the Respondent Sample

The first study involves comparing key estimates of the full sample (respondents and nonrespondents) to estimates of the respondent sample for Wave 2 of the SIPP. Using the variables available in Wave 1, differences between the weighted statistics (means or percentages) of the full sample and the respondent sample are examined. Initial weights[3], which are the weights that have incorporated unit nonresponse from Wave 1, are used in estimating the weighted statistics. Examining the differences between the estimates gives insight into the magnitude and direction of bias.

Fay's method of Balanced Repeated Replication[4] (BRR) is used to estimate the variance of the difference between the full sample and the respondent sample estimates. The resampling method accounts for both the complex survey design and the random error (variability) due to sampling and nonresponse. Fay's method is also used to estimate the covariance[5] matrix of the difference between the distribution of the full sample and nonrespondent sample.

---

[3] Using initial weights ensures that any bias observed is from Wave 2.

[4] The 116 replicate weights using the Fay's method of BRR are produced in SAS using the PROC SURVEYMEANS procedure with adjustments for sample design and the VARMETHOD=BRR(FAY=0.5 OUTWEIGHTS=) statement. Initial weights are used to produce replicate weights.

[5] For further details on estimating the covariance matrix using resampling methods, please see Efron, B. (1982) [2] and Young, P. (2010) [20].

For variables with more than two levels (e.g., age, CBSA area, education, household size, household type, region), the distribution of the full sample is compared to the distribution of respondent sample using a chi-square test to examine potential nonresponse bias [18].

Analyzing Estimates by Predicted Response Propensity Quintiles

The second study involves analyzing key estimates by predicted response propensity quintiles. A response propensity model is fit using a weighted logistic regression to obtain the predicted values or response propensities using the initial weight with adjustments for sample design. The logistic regression model predicts the probability of a household being a Wave 2 respondent as a function of Wave 1 variables, and the predicted values or response propensities obtained from the models are used to group the sample into quintiles.

The dependent variable in the logistic regression is a binary indicator defined as Wave 2 respondent or nonrespondent. The independent variables in the model are region, age, gender, race, Hispanic origin, education, and marital status of reference person from Wave 1 of the panel.

The response propensities from the logistic regression analysis are used to group the sample households into response propensity quintiles. Seven key estimates are then examined for potential nonresponse bias. Percentages and means are calculated for the key variables within each propensity group (quintile), and hypothesis testing is conducted to determine whether the statistics vary by subgroup. The low response propensity households can be treated as proxies for nonrespondents, and a two sample t-test is used to compare estimates of the low propensity group (quintile) to the remainder of the sample [1]. A difference in a statistic over increased response propensity quintiles is indicative of the risk of nonresponse bias.

Calculating the R-Indicator

The third analysis involves calculating the R-indicator to examine the potential for nonresponse bias. R-indicators measure how representative the respondents are compared to the original sample or population [19].

The estimation of the R-indicator involves fitting a weighted logistic regression model that predicts the probability of being a Wave 2 respondent as a function of Wave 1 variables to estimate response propensities. The standard deviation of the response propensities is obtained from the model, and the R indicator is estimated by the following equation [19]:

$$\hat{R} = 1 - 2S_{\hat{\rho}} = 1 - 2\sqrt{\frac{1}{\sum_{i=1}^{42,029} w_i - 1}\sum_{i=1}^{42,029} w_i\left(\hat{\rho}_i - \bar{\hat{\rho}}\right)^2},$$

where
w_i is the initial weight
$\hat{\rho}_i$ are the response propensities estimated using the logistic regression model.

Values of the R-indicator that are close to one are an indication of strong representativeness since large values occur when the standard deviation of the response propensity is small. This means that the response propensities tend to be more similar, and therefore the respondents are more likely to be representative of the sample or population. Values close to zero are an indication of weak representativeness or that the respondents are less likely to be representative of the sample or population.

A confidence interval is constructed for the estimate of the R-indicator using Fay's method of BRR, which accounts for the sample design and the model used to estimate the response propensities [19].

Results

Comparing Key Estimates of the Full Sample to Respondents-only Sample

Table 1 summarizes the results of comparing estimates of the full sample to the respondent sample for Wave 2 of the 2008 SIPP. Estimates for the full sample use the initial weight; and two estimates are given for the respondent sample, one weighted by the initial weight and the other weighted by the nonresponse adjusted weight. While results with the initial weight can identify the presence of nonresponse bias, the results from using the nonresponse adjusted weight will indicate if the weighting is helping to correct for some of the bias.

Viewing "Medicare" as an example reveals that 27.8% of the entire sample has at least one member of the household receiving Medicare coverage and 28.8% of the respondents have at least one member of the households receiving coverage when using the initial weights. When using the weights that have been adjusted for nonresponse at Wave 2, 28.3% of the respondents have households receiving Medicare coverage.

Table 1. Comparison of Estimates from Full Sample and Respondent Sample

	All Sample Cases Initial Weight		Respondents Initial Weight				Respondents Nonresponse Adjusted Weight			
Variables	Percent/ Mean	Std Error	Percent/ Mean	Std Error	Relative Difference (%)		Percent/ Mean	Std Error	Relative Difference (%)	
Medicare	27.82	0.25	28.77	0.26	3.32	*	28.32	0.25	1.79	*^
Medicaid	17.17	0.22	16.72	0.22	-2.70	*	16.99	0.23	-1.03	*^
SNAP	7.56	0.14	7.21	0.14	-4.88	*	7.34	0.14	-3.11	*^
SSI	4.57	0.11	4.56	0.12	-0.40		4.61	0.12	0.73	
Social Security	30.14	0.26	31.12	0.26	3.14	*	30.69	0.26	1.79	*^
Household Earnings	$4307.26	39.81	$4340.47	41.07	0.77	*	$4330.15	40.85	0.53	*^
Household Income	$5391.53	41.07	$5453.93	41.28	1.14	*	$5428.90	41.00	0.69	*^
Region										
Northeast	18.85	0.17	18.82	0.20	-0.12		18.81	0.20	-0.18	
Midwest	22.90	0.24	23.22	0.26	1.41	*	23.05	0.26	0.66	*^
South	35.94	0.27	35.96	0.31	0.06		36.00	0.31	0.17	
West	22.32	0.22	21.99	0.26	-1.48	*	22.14	0.27	-0.80	*^
Urban/Rural										
Rural	20.60	0.58	21.18	0.58	2.73	*	21.02	0.58	1.99	^
Urban	79.40	0.58	78.82	0.58	-0.73	*	78.98	0.58	-0.53	^
Within PSU Strata										
Non-low income strata	76.80	0.32	77.29	0.33	0.63	*	77.01	0.33	0.26	^
Low income strata	23.20	0.32	22.71	0.33	-2.14	*	22.99	0.33	-0.88	^
CBSA Type										
In Principal City of MSA	32.83	0.75	32.18	0.74	-2.03	*	32.44	0.74	-1.23	*^
In MSA not in Principal City	50.30	0.92	50.62	0.94	0.62	*	50.45	0.94	0.30	^
Not in MSA or Principal City	16.86	1.29	17.20	1.32	1.96	*	17.11	1.31	1.46	*
HH size										
1	28.63	0.27	28.18	0.29	-1.58	*	28.23	0.29	-1.42	*^
2	34.74	0.28	35.32	0.29	1.63	*	35.13	0.29	1.11	*^
3	15.08	0.17	14.95	0.18	-0.87		15.04	0.18	-0.30	
4+	21.54	0.20	21.54	0.21	-0.01		21.60	0.21	0.26	
Assets										
Bonds/Etc.	39.05	0.32	40.33	0.36	3.16	*	39.86	0.35	2.03	*^
Minimal	60.95	0.32	59.67	0.36	-2.14	*	60.14	0.35	-1.35	*^
Income Type										
Welf/Etc.	18.98	0.24	18.44	0.23	-2.92	*	18.74	0.23	-1.25	*^
Other	81.02	0.24	81.56	0.23	0.66	*	81.26	0.23	0.29	*^
Tenure										
Owner	68.02	0.28	70.33	0.29	3.28	*	69.60	0.30	2.27	*^
Renter	31.98	0.28	29.67	0.29	-7.78	*	30.40	0.30	-5.21	*^

Table 1 continues on the next page.

(Table 1 continued)

Variables	All Sample Cases Initial Weight Percent	Std Error	Respondents Initial Weight Percent	Std Error	Relative Difference (%)		Respondents Nonresponse Adjusted Weight Percent	Std Error	Relative Difference (%)	
Age of Householder										
Under 25	4.69	0.13	4.09	0.13	-14.61	*	4.18	0.13	-12.29	*
25-34	15.21	0.21	14.33	0.21	-6.13	*	14.54	0.21	-4.57	*^
35-54	39.97	0.25	40.08	0.27	0.27		40.26	0.27	0.71	*^
55+	40.13	0.27	41.50	0.29	3.29	*	41.02	0.29	2.17	*^
Gender of Householder										
Male	46.82	0.25	46.77	0.26	-0.10		46.73	0.26	-0.19	
Female	53.18	0.25	53.23	0.26	0.09		53.27	0.26	0.17	
Race of Householder										
Non-Black	88.09	0.23	88.50	0.24	0.47	*	88.20	0.25	0.12	^
Black	11.91	0.23	11.50	0.24	-3.62	*	11.80	0.25	-0.93	^
HH Type Reference person is female with no husband present and with her own children less than 16 years old	1.22	0.08	1.20	0.08	-1.42		1.21	0.08	-0.62	
Reference person is 65 years old or older	22.23	0.23	23.04	0.23	3.52	*	22.57	0.23	1.54	*^
Other	76.55	0.25	75.76	0.25	-1.05	*	76.21	0.25	-0.45	*^
Educational Attainment of Householder										
Less than High School	4.61	0.14	4.57	0.15	-0.71		4.61	0.15	0.17	
High School, no diploma	6.75	0.14	6.60	0.15	-2.23	*	6.65	0.15	-1.41	*^
High School graduate	58.64	0.32	58.38	0.35	-0.44	*	58.44	0.35	-0.35	*^
College graduate	30.00	0.32	30.44	0.32	1.44	*	30.29	0.32	0.95	*^

Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel. For information on sampling and nonsampling error see <http://www.census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements.html>.

*The estimate for respondents is significantly different from the full sample estimate at the 90% confidence level.

^The nonresponse adjustment helped to reduce the bias. The relative difference for the nonresponse adjusted estimate is either no longer significant (relative difference was significant for the initial weighted estimate) or the difference is smaller than the relative difference using only the initial weight at the 90% confidence level.

Relative differences between the estimates are displayed in Table 1. Using the initial weight, mean household earnings and mean household income are both overestimated by about 1% in the final sample. The percent of households with at least one household member receiving Medicare coverage is overestimated by 3%, and the percent receiving Social Security is overestimated by 3%. The percent of households with at least one household member receiving Medicaid coverage is underestimated by 3%, and the percent receiving SNAP benefits is underestimated by 5%. The percent of households receiving SSI is neither overestimated nor underestimated.

With the exception of the age and tenure variables, none of the relative differences is greater than 5% when using the initial weight. The percent of households in rural areas are overestimated, while the percent of renters, percent of householders who are less than 35 years old, percent of households in low income strata, percent located within MSA in principal city, and percent of blacks are underestimated. The percent of male householders is neither overestimated nor underestimated.

Differences between the relative differences calculated using the initial weight and nonresponse adjusted weight reveal whether the adjustment helped reduce nonresponse bias. Using the nonresponse adjusted weight, the relative difference drops for almost all of the estimates. None of the relative differences using the nonresponse adjusted weight is greater than 3% with the exception of the age and tenure variables.

The distribution of the estimates with more than two categories was tested using the chi-square test. The distribution of the full sample was significantly different from the respondents for CBSA Type, household size, age of householder, household type, and education.

Analyzing Estimates by Predicted Response Propensity Quintiles

Table 2 presents results of the logistic regression model predicting the probability of a household being a Wave 2 respondent as a function of Wave 1 variables. Significant effects are observed for region, age, race, gender, education, and marital status of reference person. The Midwest has 20% greater odds of responding than the Northeast. Sample units with householders who are less than 45 years of age have lower odds and units with householders who are greater than 55 years of age have higher odds of being interviewed than those who are between the ages of 45 and 54. Households with female householders have 7% greater odds of responding than households with male householders. Compared to sample units with a White householder, units with a Black or Asian householder have lower odds of being interviewed. Households that have a householder with no high school diploma have lower odds of responding, and households with a householder with a bachelor's or a higher degree have higher odds of responding than households with a householder whose highest level of educational attainment is high school diploma or associate degree. Households with a reference person who is widowed, divorced, separated, or never married have lower odds of responding than households with a married householder.

Table 2. Odds Ratios from Logistic Regression of Wave 2 Response

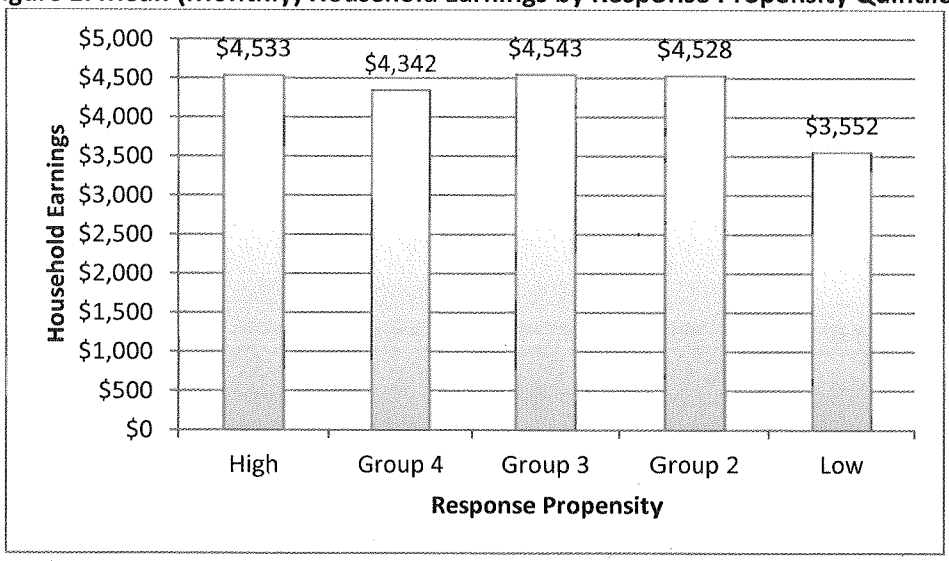
Model Variables	Odds Ratio	90% CI
Region		
Northeast†	1.00	--
Midwest	1.19*	(1.07, 1.33)
South	1.07	(0.98, 1.16)
West	0.95	(0.86, 1.05)
Age		
15-24	0.42*	(0.38, 0.47)
25-34	0.57*	(0.52, 0.62)
35-44	0.77*	(0.7, 0.85)
45-54†	1.00	--
55-64	1.17*	(1.06, 1.29)
65+	1.32*	(1.19, 1.47)
Gender		
Male†	1.00	--
Female	1.07*	(1.01, 1.13)
Race		
White†	1.00	--
Black	0.79*	(0.73, 0.85)
Asian	0.68*	(0.58, 0.79)
Other	0.88	(0.77, 1.01)
Hispanic Origin		
Hispanic origin†	1.00	--
Non-Hispanic origin	1.08	(0.98, 1.19)
Education		
Less than high school	0.92	(0.81, 1.04)
9th-12th grade, No diploma	0.89*	(0.81, 0.98)
High school graduate/Associate degree†	1.00	--
Bachelor's or higher degree	1.20*	(1.12, 1.28)
Marital Status		
Married†	1.00	--
Widowed	0.72*	(0.64, 0.82)
Divorced	0.78*	(0.72, 0.84)
Separated	0.55*	(0.48, 0.64)
Never married	0.71*	(0.66, 0.77)

Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.
† Reference category
* P-value < 0.10

Predicted values or response propensities from the logistic regression analysis are used to group the sample households into response propensity quintiles. Figures 1-7 show the estimated means or percent for income, earnings, Medicaid coverage, Medicare coverage, Social Security receipt, SSI receipt, and SNAP receipt within each predicted response propensity quintile from fitting the logistic regression model.

Figure 1 presents the mean monthly household earnings in Wave 1 of the 2008 SIPP by response propensity quintiles. If the households in the "low" propensity quintile are treated as proxies for nonrespondents [1], the mean household earnings may be overestimated in the sample since the estimate of $3,552 for the "low" response propensity quintile is significantly less than the estimate of $4,487 for the remainder of the sample (one-sided t-test p-value: <0.001).

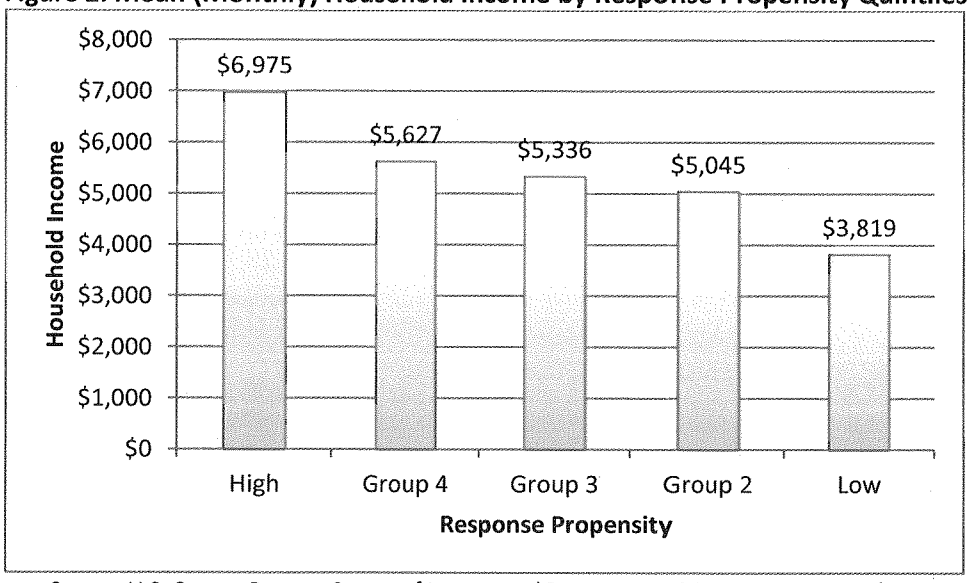
Figure 1. Mean (Monthly) Household Earnings by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel

Figure 2 presents the mean monthly household income by response propensity quintiles. If the households in the "low" propensity quintile are treated as proxies for nonrespondents, the results suggest that the mean household income may be overestimated in the sample. The estimate of $3,819 for the "low" response propensity quintile is significantly less than the estimate of $5,746 for the remainder of the sample (one-sided t-test p-value: <0.001).

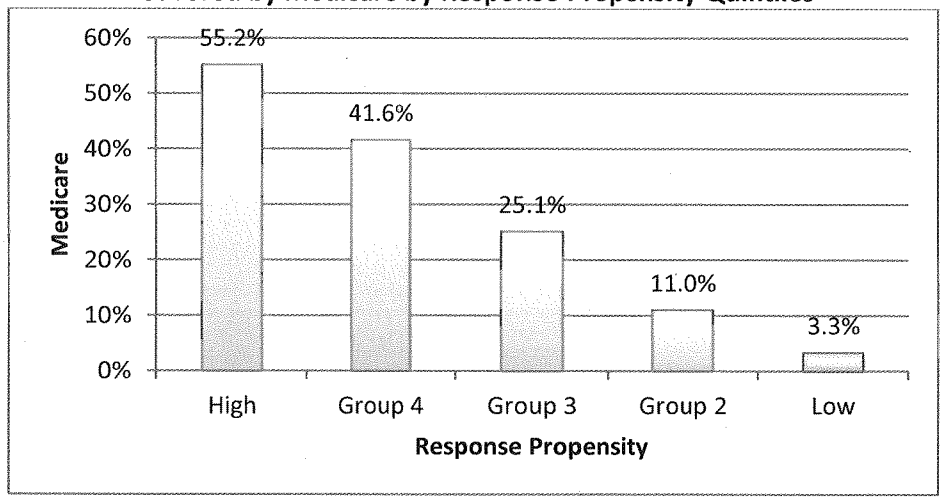
Figure 2. Mean (Monthly) Household Income by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.

Figure 3 presents the percent of households that was covered by Medicare by response propensity quintiles. The estimate of 3.3% for the "low" response propensity quintile is significantly less than the estimate of 33.6% for the remainder of the sample. The results suggest that the percent of households covered by Medicare may be overestimated in the sample (one-sided t-test p-value: <0.001).

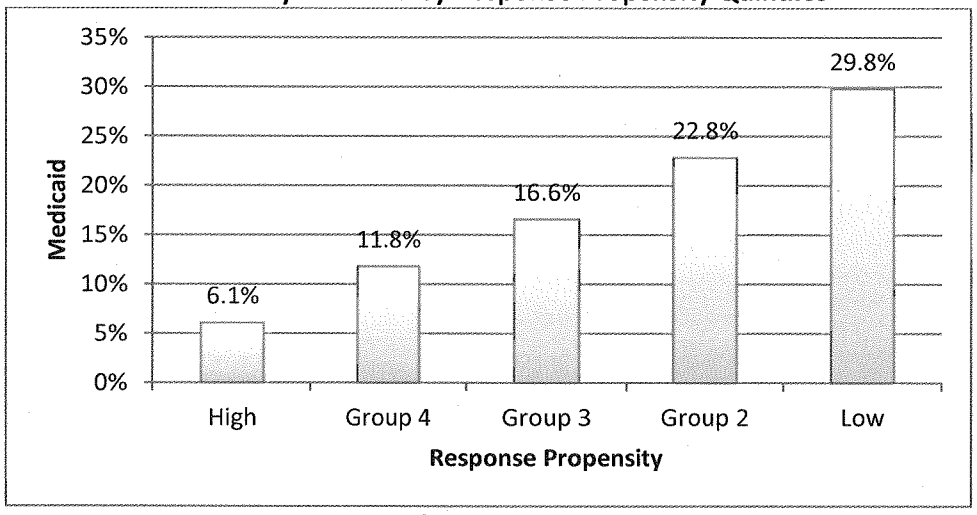
Figure 3. Percent of Households where at least One Household Member was Covered by Medicare by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.

Figure 4 presents the percent of households that was covered by Medicaid by response propensity quintiles. If households in the "low" response propensity quintile are treated as proxies for nonrespondents, their estimate of 29.8% is significantly greater than the estimate of 14.2%. The results suggest that the percent of households covered by Medicaid may be underestimated in the sample (one-sided t-test p-value: <0.001).

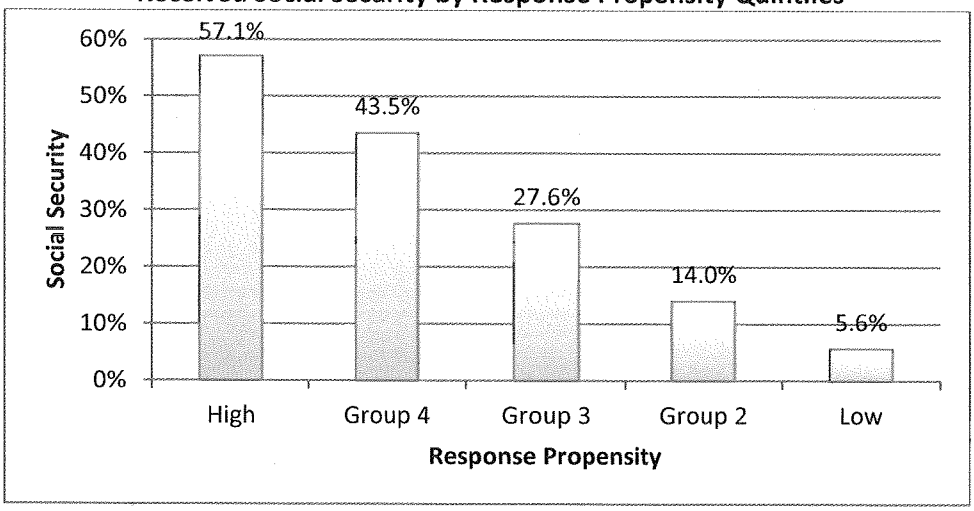
Figure 4. Percent of Households where at least One Household Member was Covered by Medicaid by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.

Figure 5 presents the percent of households that received Social Security by response propensity quintiles. The estimate of 5.6% for the "low" response propensity quintile is significantly less than the estimate of 36.0% for the remainder of the sample. The results suggest that the percent of households that received Social Security may be overestimated in the sample (one-sided t-test p-value: <0.001).

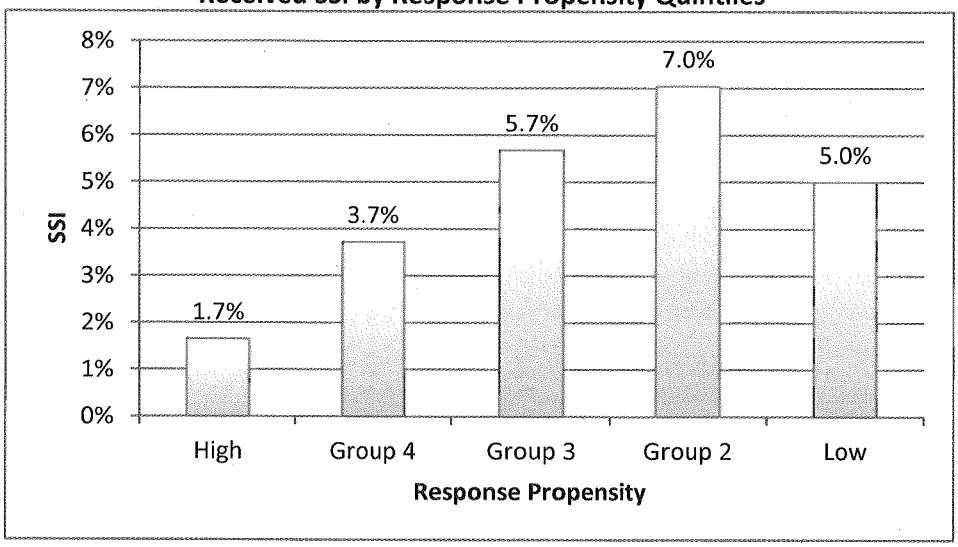
Figure 5. Percent of households where at least One Household Member Received Social Security by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.

Figure 6 presents the percent of households that received SSI by response propensity quintiles. The estimate of 5.0% for the "low" response propensity quintile is significantly greater than the estimate of 4.5% for the remainder of the sample. If households in the "low" response propensity quintile are treated as proxies for nonrespondents, the results suggest that the percent of households that received SSI is underestimated in the sample (one-sided t-test p-value: 0.03).

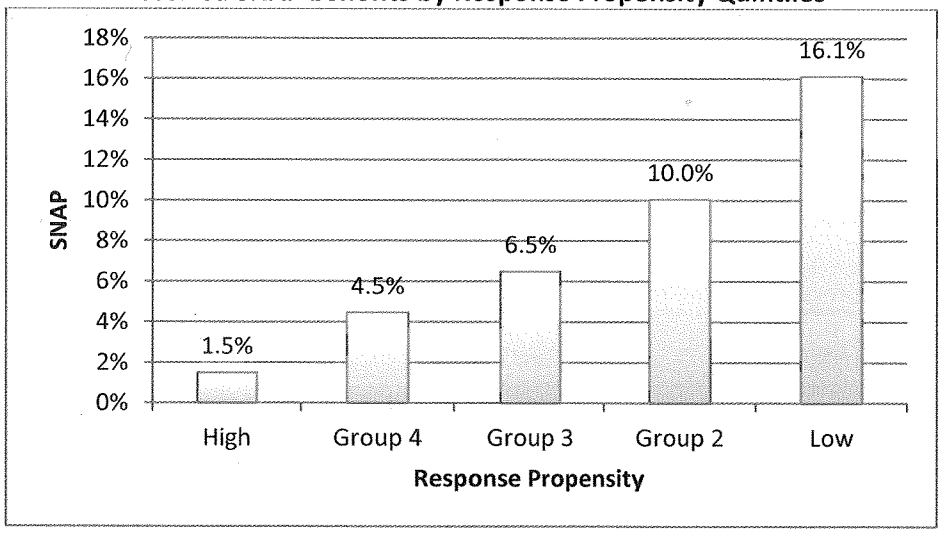
Figure 6. Percent of households where at least One Household Member Received SSI by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.

Figure 7 presents the percent of households that received SNAP benefits by response propensity quintiles. The estimate of 16.1% for the "low" response propensity quintile is significantly greater than the estimate of 5.5% for the remainder of the sample. The results suggest that the percent of households that received SNAP benefits may be underestimated in the sample (one-sided t-test p-value: <0.001).

Figure 7. Percent of households where at least One Household Member Received SNAP benefits by Response Propensity Quintiles



Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel.

R-Indicator

Table 3 lists the set of independent variables for two logistic regression models predicting the probability of a household being a Wave 2 respondent. The variables included in Model 1 are the variables that have been included in the model for determining the response propensity quintiles. The variables for Model 2 are the nonresponse weighting adjustment variables.

Table 3. Model 1 and Model 2 Variables in Logistic Regression of Wave 2 Response

Model 1
Age of Reference Person (Six levels: 15-24, 25-34, 35-44, 45-54, 55-64, 65+)
Education of Reference Person (Four levels: Less than high school; 9th-12th, no diploma; High school graduate; Bachelor's or higher)
Gender of Reference Person (Two levels: Male, Female)
Hispanic Origin of Reference Person (Two levels: Hispanic origin, Non-Hispanic origin
Marital Status of Reference Person (Five levels: Married, Widowed, Divorced, Separated, Never married)
Race of Reference Person (Four levels: White, Black, Asian, Other)
Region (Four levels: Northeast, Midwest, South, West)
Model 2
Age of Reference Person (Four levels: Under 25, 25-34, 35-54, 55+)
Assets (Two levels: Bonds/Etc, Minimal)
CBSA area (Three levels: In Principal City of MSA, In MSA not in Principal City, Not in MSA or Principal city)
Education of reference person (Four levels: Less than high school; 9th-12th, no diploma; High school graduate; Bachelor's or higher)
Gender of Reference Person (Two levels: Male, Female)
Household Size (Four levels: 1 persons in the household, 2 persons in the household, 3 persons in the household, 4+ persons in the household)
Household Type (Three levels: Ref. person is female with no husband present and with her own children less than 16 years old, Ref. is 65 years old or older, Other)
Income Type (Two levels: Welf/Etc - Household received income from SSI, AFDC, other welfare, WIC, Food Stamps, or Medicaid; Other)
Race of Reference Person (Two levels: Black, Non-Black)
Region (Four levels: Northeast, Midwest, South, West)
Tenure (Two levels: Renter, Owner)
Urban/Rural (Two levels: Urban area, Rural area)
Within PSU Strata (Two levels: Low income strata, Non-low income strata)

The estimate of the R-indicator obtained for the two models is 0.9087 for Model 1 and 0.8909 for Model 2. The large estimate of the R-indicator for both models suggests that there is a high likelihood the respondent sample is representative of the full sample.

Conclusion

This analysis used three methods to examine the potential for nonresponse bias, including comparison of full sample and respondent sample estimates, analyzing estimates by predicted response propensity quintiles, and calculating R-indicators. Since each method has weaknesses, no single method is sufficient on its own to draw accurate conclusions. By incorporating multiple approaches in the analysis, results were compared across methods, leading to clearer and stronger conclusions.

Some areas of potential bias were identified in the results. Wave 2 of the SIPP may be overestimating household earnings, household income, and participation in Medicare and Social Security. Participation in Medicaid and SNAP may be underestimated in the responding sample. However, none of the relative differences for these key measures were greater than 5%. There does not appear to be nonresponse bias associated with the percent of households receiving SSI.

Tenure, age, and race displayed the largest relative differences among demographic and noninterview weighting adjustment variables. Renters, households with Black householders, and households with reference persons less than 35 years of age may be underestimated in the responding sample, while households with householders who are 55 and over may be overestimated.

The logistic regression models that were fit to determine the propensity quintiles and R-indicators pointed to similar areas of potential bias due to certain subgroups' likelihood of responding. The estimate of the R-indicator differed depending on which covariates were included in the model. However, the large estimate of the R-indicator for both models suggested that there is a high likelihood the respondent sample is representative of the full sample or population.

A limitation of this research is that the majority of nonresponse for the SIPP happens in the first wave. Therefore, the examination of characteristics of respondents and nonrespondents in later waves may not give accurate insights into the major component of the nonresponse bias for the key estimates. In addition, results from the propensity quintiles and R-indicators are highly dependent on model assumptions and parameterization.

As mentioned in the Results section, many of the significant differences in the comparison of estimates are likely due to the large sample size of the survey. In classical statistics, large sample sizes lead to an increase in the probability of rejecting the null hypothesis. However, the relative differences and the R-indicators reveal that the significant differences observed are not a cause for concern and do not directly imply that nonresponse bias exists.

There appears to be evidence of potential bias due to nonresponse for some statistics in Wave 2 of the SIPP based on the analyses conducted. However, the biases are significantly reduced with the use of noninterview adjustments and may be further reduced with second stage adjustments. Despite the significant differences observed between compared estimates, the statistical significance of the differences does not imply practical significance since the small relative differences and the large estimates of the R-indicators confirm that the respondent sample is highly representative of the sample or population.

Similar analyses of the cross sectional estimates for later waves of the SIPP 2008 Panel can be conducted, as well as, investigating the nonresponse bias in the longitudinal estimates. Another relevant analysis is currently in progress which uses IRS income data to measure nonresponse bias in the SIPP. The results of this study are expected by the end of the year.

References

1. Dahlhamer, J.M. & Simile, C.M. (2009). Subunit Nonresponse in the National Health Interview Survey (NHIS): An Exploration Using Paradata. Paper presented at *Proceedings of the Government Statistics Section of the American Statistical Association.*
2. Efron, B. (1982). "The jackknife, the bootstrap, and other resampling plans". *Society of Industrial and Applied Mathematics CBMS-NSF Monographs*, 38.
3. Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675.
4. Mack, S. & Petroni, R. (1994). Overview of SIPP Nonresponse Research. Paper presented at the *Proceedings for Fifth International Workshop on Household Survey Non-response*, September 26-28.
5. Memorandum from Ruth Ann Killion to Jason Fields. "Nonresponse Bias Analysis for Wave 1 of the 2008 Survey of Income and Program Participation (SIPP) (ALYS-12)", June 03, 2013.
6. Memorandum from Alan R. Tupek to Chester E. Bowie. "SIPP 2004: Cross-Sectional Weighting Specifications for the Second and Subsequent Waves (WGT-5)", May 27, 2004.
7. Office of Management and Budget Standards and Guidelines for Statistical Surveys. http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf.
8. Presentation of "Practical Tools for Nonresponse Bias Studies". JPSM Short Course presented by Jill M. Montaquila and Kristen M. Olson on September 30, 2011.
9. Sae-Ung, S., Sissel, C. D., & Mattingly, T. L. (2007). Analytical Comparison of the SIPP and CPS_ASEC Key Longitudinal Estimates. Paper presented at the *Federal Committee on Statistical Methodology Research Conference*, November 5-7.
10. Sawyer A. & Perer P. (1983). The Significance of Statistical Significance Tests in Marketing Research. *Journal of Marketing Research.* Vol. XX, pp. 122 – 133
11. Schouten B., Cobben F., & Bethlehem J. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, 35,101-113.
12. Survey of Income and Program Participation Quality Profile (1998). Available at http://www.census.gov/sipp/workpapr/wp230.pdf.
13. Survey of Income and Program Participation Users' Guide. Available at http://www.census.gov/programs-surveys/sipp/methodology/users-guide.html.
14. U.S. Census Bureau Memorandum to Nonresponse Oversight Team on March 21, 1995 from Nonresponse Research Team. Preliminary Final Report on SIPP Nonresponse.
15. U.S. Census Bureau Statistical Quality Standards http://cww.census.gov/msdir/docs/Quality_Standards_cww.pdf.
16. U.S. Census Bureau Memorandum (2002), *SIPP 2000 Redesign: Creating Variance Codes for Public Use Files.*
17. U.S. Census Bureau Memorandum (2012), *Survey of Income and Program Participation (SIPP) 2008 Panel: Source and Accuracy Statement for Wave 1 to Wave 11 (core) Public Use Files.* Available at http://www.census.gov/content/dam/Census/programs-surveys/sipp/tech-documentation/source-accuracy-statements/2008/SIPP%202008%20Panel%20Wave%2005%20-%20Core%20Source%20and%20Accuracy%20Statements.pdf.

18. van de Kerckhove, W., Krenze T., & Mohadjer L. (2006). Approaches to a Nonrespones Bias Analysis in an Adult Literacy Survey. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 3790-3795.
19. Witt, M. B. (2010). Estimating the R-indicator, Its Standard Error and Other Related Statistics with SAS and SUDAAN. Paper presented at *JSM Proceedings, Section on Survey Research Methods. American Statistical Association.*
20. Young, P. (2010). "Jackknife and Bootstrap resampling methods in statistical analysis to correct for bias," lecture notes, Physics Department, University of California – Santa Cruz, Santa Cruz, CA.