Do Interviewer Observations Agree

with Previously Collected Survey Data?

Dan Perez-Lopez*
U.S. Census Bureau
SEHSD Working Paper Number 2019-26

Prepared for
American Association for Public Opinion Research (AAPOR)
Annual Conference
May 2019
Toronto, Ontario, Canada

Abstract

In recent years, the reduction in response rates to large government surveys has led to an increased importance in measuring nonresponse bias.[1] Meanwhile, statistical agencies have been under pressure to spend limited resources efficiently. This has led to interest in the use of interviewer-generated paradata to analyze nonresponse bias. It has also led to an interest in using administrative or previously collected survey data as part of the data analysis process. This methodological brief uses survey data linked to previously collected household data to assess the agreement between interviewer-generated observations and previously collected data for nonrespondents. As part of this evaluation, we explore whether there are differences in agreement by the type of nonrespondent household, and the impact of substituting this previously collected data for variables which are currently imputed.

In this study, noninterviewed households are classified as those contacted by interviewers, not contacted by interviewers, or other. This analysis suggests that agreement between interviewer-generated race for noninterviews and previously collected household data is higher for contacted households than for noncontacted and residual households. It finds that agreement between interviewer-generated sex for noninterviews and previously collected household data is not different across the three noninterview types. When previously collected data was substituted for imputed interviewer-generated race the resulting distributions did not differ significantly. This suggests that previously collected household data may be able to serve a role in the analysis of nonresponse bias, which is a topic for future work.

---

[1] De Leeuw ED and De Heer W (2002). "Trends in household survey nonresponse: a longitudinal and international comparison." Groves RM, Dillman DA, Eltinge JL and Little RJA (eds) Survey Nonresponse. New York: John Wiley, 41–54.

Do Interviewer Observations Agree with Previously Collected Survey Data?

Introduction

The Survey of Income and Program Participation (SIPP) is a household survey conducted by computer-assisted personal interviewing (CAPI) by the US Census Bureau. In the case of interviews that end as incomplete, field interviewers record four items for use in nonresponse bias analysis: the gender and race of the householder, the number of household members, and whether the unit is owned or rented. This brief uses 2014 Wave 1 SIPP data linked to the Master Demographic pilot program to assess the agreement between field interviewer-generated paradata and information collected previously in household surveys and Censuses.

The Master Demographics pilot program (MD) at the US Census Bureau links data collected in the 2000 and 2010 Decennial Censuses with data collected from multiple years of the American Community Survey, Current Population Survey, SIPP, and other household surveys, from 2000 to the present. These data are matched at the address and person ID-level, by a protected identification key, or PIK. This allows information collected in previous surveys to be available for respondents and nonrespondents in an ongoing household sample, like the SIPP, in cases where a match can be made by address and PIK to people in sampled units.

Wave 1 response rates for the SIPP have been declining over the last few panels. In the 2001 panel, the Wave 1 response rate for was 87 percent. In the 2004 panel, about 85 percent of all eligible housing units participated in the Wave 1 interview. In the 2008 panel, the weighted response rate for the Wave 1 interview was 81 percent.[2] In the 2014 panel, the SIPP began conducting one annual interview instead of three interviews a year. This change required re-engineering the interview, and the new instrument resulted in longer, but less frequent, interviews. In the 2014 panel, the weighted response rate for the Wave 1 interview was 69 percent.

The Office of Management and Budget (OMB) requires survey programs to conduct a nonresponse bias analysis if unit response rates fall below 80 percent. In order to conduct this nonresponse bias analysis, SIPP field interviewers are asked to record four items for noninterviewed eligible housing units: the sex and race of the householder, the number of household members, and whether the unit is owned or rented. Determining whether low response rates are associated with nonresponse bias requires analysts to determine if respondents and nonrespondents differ on survey variables of interest. In the nonresponse bias analysis conducted by the SIPP, all four noninterview frame variables were correlated with survey variables of interest. Given that sex and race of householder are correlated with the survey variables of interest, it is worthwhile to explore how these variables are generated by interviewers.[3] The motivation for this study began with the question, "How do interviewers generate the frame noninterview variables for sex and race that are used in the SIPP nonresponse bias analysis?"

---

[2] For W1 response rates for the 2001 and 2004 panels, see their respective Preliminary Wave 1 Core Microdata File. For the 2008 W1 panel response rate, see McMillan, S and Culver III, R, "Nonresponse Bias Analysis for Wave 1 of the 2008 Survey of Income and Program Participation (SIPP)," Technical Report, 2013. For 2014 W1 panel response rate, see Westra, A and Nwaoha-Brown, F "Nonresponse Bias Analysis for Wave 1 of the Survey of Income and Program Participation (SIPP)," Technical Report, 2017.

[3] See Westra and Nwaoha-Brown (2017), for a detailed description of the SIPP nonresponse bias analysis and the frame variables that enter into it.

Another question was whether the ability of field interviewers to determine sex and race of noninterviewed people differed by the type of noninterview. To wit, was it easier for field interviewers to determine sex and race of householders who they spoke with and who refused the interview than with householders they were unable to make contact with? The processes field interviewers use to generate these variables for noninterviewed households are not well-understood. Meanwhile, not all interviewers report sex and race for noninterviewed households, and there is little knowledge of the process by which they decide to report or to fail to report these variables. In the case of the SIPP, whether these vary by the type of noninterview has not been thoroughly examined.

Data

Survey of Income and Program Participation (SIPP)

The Survey of Income and Program Participation (SIPP) is a household survey conducted through computer-assisted personal interviewing (CAPI) by the US Census Bureau. In the 2014 panel, sampled households were interviewed once a year for 4 years, with a one-year reference period. As shown in Table 1, in Wave 1 of the 2014 panel, referencing calendar year 2013, approximately 53,000 units were sampled, but after interviewers began data collection, only approximately 42,500 of these units were found to be eligible units.[4] Out of these, 30,000 units either completed an interview, or completed a sufficient partial interview, for an unweighted response rate of 70.1 percent. The remaining 12,500 units were classified as noninterviews, and a subset of these units is the focus of this research.

As part of the data collection process, SIPP interviewers are instructed to make multiple attempts in person and by phone to schedule and complete an interview. Noninterviewed units are divided into three main groups based on the reason why an interview could not be obtained: Type A, Type B, or Type C. If an occupied unit results in a noninterview, the unit is classified as Type A noninterview. When field interviewers determine that a sample unit is unoccupied, or no eligible persons are living at the address, the unit is coded as a Type B noninterview. When a field interviewer determines that a sampled address does not exist, or is unable to be occupied, the case is coded as a Type C noninterview.

Among the remaining housing units eligible to be interviewed, the SIPP interviews all adults in a sampled housing unit, and allows interview by proxy. A subset of items is collected, by proxy, for children under the age of 15. At the end of the data collection period, some units have completed the entire SIPP interview, or have completed at least the demographic questions and the event history calendar questions.[5] These units are coded as either completes or complete partials, respectively.

---

[4] All figures in this report are rounded to meet Census guidance on disclosure.

[5] An Event History Calendar is a tool for reducing error due to recall in data collection. For a description of the event history calendar used in the SIPP, see the National Academies of Sciences, Engineering, and Medicine. (2018). *The 2014 Redesign of the Survey of Income and Program Participation: An Assessment.* Washington, DC: The National Academies Press.

Table 1: Interview Outcomes for Sampled Units, SIPP 2014 W1 (unweighted)

Outcome	N	Percent	Std Error of Percent	MoE 90% CI
Complete Interview	27,000	50.9	0.2	0.4
Complete Partial Interview	2,500	4.7	0.1	0.2
NonInterview Type A	12,500	23.6	0.2	0.3
NonInterview Type B	8,500	16.0	0.2	0.3
NonInterview Type C	2,200	4.2	0.1	0.1
TOTAL SAMPLED UNITS	53,000	100		

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1.

After the Type B, Type C, complete and complete partials are accounted for, the remaining sampled units are designated as Type A household-level noninterviews. This includes interviews that were never obtained as well as interviews that were begun but did not reach the complete partial threshold. Type A noninterviews are cases where a sample unit is occupied by persons presumed to be eligible for interview, but neither a complete or partial interview is obtained. This study focuses on noninterviewed housing units coded as Type A.

There are six outcome codes available to field interviewers for Type A noninterviews as shown in Table 2. For the purposes of this study, we've chosen to combine these classifications into three categories.

Table 2: Type A Noninterview Outcomes

Contact	NonContact	Other
Language Problem	Unable to Locate	Other Occupied (specify)
Household Refused	No One Home (NOH)	
	Temporarily Absent (TA)	

Source: "Source and Accuracy Statement for the 2014 Panel Wave 1 Public Use Files of the Survey of Income and Program Participation (SIPP)", Technical Report, 2017.

"Language problem" and "household refused" both imply that the field interviewer spoke with someone in the household, so it is logical to combine those two categories. In these cases, the interviewer made contact with someone in the household, so we'll categorize these as **Contact**. "Unable to locate," "No one home," and "temporarily absent" all imply that the interviewer was unable to speak to anyone in the household, so it is sensible to combine these three categories and call them **Non Contact**. The "other" occupied category may include noninterviews where household members were spoken to, and noninterviews where no household members were spoken to, so we'll consider those cases separately, and call them **Other**. After this reclassification, Table 3 shows the number of noninterviews by type.

Table 3: Number and Type of Type A Noninterviews for Analysis, SIPP 2014 W1 (unweighted)

Outcome	N	Percent of all Type A Noninterviews	Std Error of Percent	MoE 90% CI
Language Problem or Household Refused (**Contact**)	10,000	79.7	0.4	0.6
Unable to Locate, No One Home, or Temporarily Absent (**NonContact**)	1,600	12.7	0.3	0.5
Other occupied	950	7.6	0.2	0.4
TOTAL	12,550	100		

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1.

As mentioned above, in order to analyze bias due to nonresponse, SIPP field interviewers are asked to record noninterview variables (sex and race of the householder, the number of household members, and whether the unit is owned or rented). However, not every field interviewer completes these items in every noninterviewed case, and they are imputed in instances when field interviewers do not record them. Moreover, these noninterview variables are reported by interviewers at different rates. To assess agreement by interviewers with previously collected survey responses, we will restrict our analysis to reported interviewer-generated noninterview variables. The rates of reporting for noninterviewed cases are shown in Table 4.

Table 4: Share of Noninterviewed Units Assigned a Noninterview Value, (unweighted) N=12,500

	N reported	Noninterview paradata, Percent reported	Std Error of Percent	MoE 90% CI
Sex	11,500	92.0	0.2	0.4
Race	11,000	88.0	0.3	0.5
Household Size	11,000	88.0	0.3	0.5
Own/Rent	11,000	88.0	0.3	0.5

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1.

Master Demographics Pilot Program (MD)

The Master Demographics pilot program (MD) at the U.S. Census Bureau links data collected in the 2000 and 2010 Decennial Censuses with data collected from multiple years of the American Community Survey (ACS), Current Population Survey (CPS), SIPP, American Housing Survey (AHS) and the National Crime Victimization Survey (NCVS) between the year 2000 to present. These data are matched at the person level by a Protected Identifier Key, or PIK. The PIK is a unique identifier, assigned by the Census Bureau to survey respondents and is constant for each person over time and across censuses and surveys.

The Master Address File (MAF) is an annually updated inventory of all known living quarters in the United States, Puerto Rico, and associated island areas. The MAF Auxiliary Reference File (MAF-ARF) links PIKs to address identifiers (MAFIDs) using Census survey data and federal administrative data.

The SIPP samples housing units, so the MAF-ARF provides a way of linking addresses to PIKs. This allows sampled housing units in the SIPP to be matched to PIKs, and then to Master Demographics. Master Demographics only includes reported values from previously collected surveys and censuses. Imputed data is not included in MD. Where address and PIK matches are made, Master Demographics allows for data to be available for both respondents and nonrespondents of the 2014 SIPP.

The number of variables available on the MD frame is limited, as variables must be harmonized over different time periods and from different surveys and censuses. New variables are added to the MD frame periodically, as resources allow. The MD data used in this analysis includes several demographic variables, including sex and race of the householder for all matched units. However, it does not include number of persons in the unit, nor tenure, defined as whether the respondent owns or rents the housing unit.

Methods

In this research, we use the MAF-ARF to link addresses sampled in Wave 1 of the 2014 SIPP panel to PIKs in the Master Demographics pilot program. As mentioned above, the SIPP 2014W1 sample included 42,500 eligible units. As shown below, some 35,500 of these units, or 83.5 percent, were linked to Master Demographics. Of these, 10,500 were noninterviews.

Table 5: Match Rate for Eligible Housing Units between SIPP and MD files, Completed and Noninterviews (unweighted)

	All eligible units		Completed Interviews		Noninterviews	
	N*	Percent	N*	Percent	N*	Percent
PIK	35,500	83.5	25,000	84.2	10,500	82.0
No PIK	7,000	16.5	4,700	15.8	2,300	18.0
All eligible units	42,500	100.0	29,700	100	12,800	100

*Due to rounding, these column sums do not agree with those in Table 1.

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1; U.S. Census Bureau Master Demographics Pilot Program, 2000-2015 internal data abstract.

These 12,800 noninterview cases required interviewers to generate values for the sex and race of the householder, the number of people in the sampled unit, and tenure. For the 82.0 percent of the 10,500 noninterviews that were linked to MD, the MD frame provides values for sex and race of the householder that can be compared to values reported by interviewers. However, interviewers did not record values for sex and race for every one of these noninterviews. In Table 6, we see that interviewers generated a value for noninterview sex about 90.5 percent of the time, and for noninterview race about 88.6 percent of the time.

Table 6: Share of Noninterviewed Units Assigned a Noninterview Value, Conditional on Unit in both SIPP and MD, (unweighted) N=10,500

	N	Percent	Std Error of Percent	MoE 90% CI
Sex	9,500	90.5	0.3	0.5
Race	9,300	88.6	0.3	0.5

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1; U.S. Census Bureau Master Demographics Pilot Program, 2000-2015 internal data abstract.

Results

If we consider these cases where interviewers provided paradata for noninterviewed units, and these units were linked to the MD file, we can examine the rate of agreement between interviewer-generated sex and race and the sex and race of the householder that is available on the MD frame. Table 7 shows rates of agreement for sex of householder for all noninterview types by noninterview classification (contact, noncontact, and other).

Table 7: Agreement between Reported Interviewer-generated Noninterview and MD for Sex (unweighted)

Noninterview Type	MD % Male	Interviewer-generated % Male	Percent Agree	Std Error	MoE 90% CI	Total N
Contact	48.5	49.7	59.8	0.5	0.9	8,000
NonContact	46.1	52.3	61.0	1.6	2.7	900
Other	47.7	49.2	62.2	2.0	3.3	600
ALL TYPES	48.2	49.9	60.0	0.5	0.8	9,500

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1; U.S. Census Bureau Master Demographics Pilot Program, 2000-2015 internal data abstract.

In looking at the agreement of interviewer-generated sex and the sex of householder available from the Master Demographic pilot program, we see that these agreed around 60 percent of the time, overall. If we look at the different kinds of noninterview outcomes, we see that there was no statistically significant difference observed in the share of agreement across the three noninterview categories, noncontact, contact, and other.

In looking at the agreement of interviewer-generated race and the race of householder available from the Master Demographic pilot program, in Table 8, we can see that agreement was much higher than in the case of sex. Overall agreement was near 95 percent. If we look at the rate of agreement by different kinds of noninterview outcomes, we see that there was a statistically significant difference observed in the share of agreement across the three noninterview categories, noncontact, contact, and other. Agreement in contacted households was over 4 percentage points higher than agreement in the

noncontacted and other households. The rates of agreement in the noncontacted and other units were not statistically significantly different at the 90 percent confidence level.

Table 8: Agreement between Reported Interviewer-generated Noninterview and MD for Race (unweighted)

Noninterview Type	MD % Nonblack	Interviewer-generated % Nonblack	Percent Agree	Std Error	MoE 90% CI	Total N
Contact	88.5	89.1	95.6	0.2	0.381	7,900
NonContact	80.8	83.6	91.1	1.0	1.630	800
Other	83.2	85.2	90.6	1.2	1.998	600
ALL TYPES	87.5	88.3	94.9	0.2	0.377	9,300

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1; U.S. Census Bureau Master Demographics Pilot Program, 2000-2015 internal data abstract.

This close agreement between interviewer-generated noninterview variables and those available from MD for race suggests a possible use for the MD data. As shown in Table 4, 88.8 percent of noninterviewed cases matched to the MD frame had a reported race variable generated by the field interviewer. However, the remaining 10.2 percent of cases had a value imputed during the data editing process. It would be possible to substitute a value from MD for those imputed cases. If we retain interviewer-generated race variables, but swap in MD values for missing data, the distribution of nonrespondent households by race shifts as shown in Table 9.

Table 9: Comparison of Edited NonInterview Race Variable with Imputations and with MD data (unweighted) N=10,500

Noninterview Variable	Nonblack %	Std Error	MoE 90% CI
Edited, with Imputations	87.9	0.3	0.5
Edited, MD swapped in for imputations	87.7	0.3	0.5

Source: U.S. Census Bureau, Survey of Income and Program Participation (SIPP) 2014 Panel, Wave 1; U.S. Census Bureau Master Demographics Pilot Program, 2000-2015 internal data abstract.

It may be worthwhile to explore using available MD data in cases where interviewer-generated paradata is unavailable.

Future Work

We have several extensions to this work that we wish to explore. First, the MD team is working to harmonize and make available a tenure variable. Once that variable is available, we could examine the agreement of interviewer-generated tenure from Master Demographics. The other interviewer-generated frame variable is the number of people living in the unit. Although the MD team has no plan to harmonize a household size variable, there is household size information available for previous respondents to the ACS. That would allow comparison of the field interviewer-generated paradata with previous ACS responses in cases where noninterviewed housing units are matched to the Master Demographics pilot program.

Once MD-level substitutes for all four interviewer-generated noninterview variables are available, we can reconstruct the nonresponse bias analysis and see if it differs from the currently constructed analysis.

A final area of interest is in how agreement between interviewer-generated frame variables and previous reports in MD varies by the age of the linked MD report. In this analysis, we've treated all reports from MD as equally likely to be accurate. However, it may be the case that more recent reports from MD are more likely to agree with interviewer-generated reports than older ones.

Bibliography

De Leeuw, ED and De Heer, W (2002). Trends in household survey nonresponse: a longitudinal and international comparison." Groves RM, Dillman DA, Eltinge JL and Little RJA (eds) Survey Nonresponse. New York: John Wiley, 41–54.

Dorinski, S and Heggeness, M, "Integrating Survey Respondents: Leveraging Existing Household Survey Data to Inform Decision-Making", Census Working Paper, 2018, available at https://www.census.gov/library/working-papers/2018/demo/SEHSD-WP2018-18.html

Groves, RM. "Nonresponse Rates and Nonresponse Bias in Household Surveys" Public Opinion Quarterly. 2006; 70(4):646–75.

Groves, RM, Couper, M, Presser, S, Singer, E, Tourangeau, R, Piani Acosta, G, Nelson, L. "Experiments in Producing Nonreponse Bias" Public Opinion Quarterly. 2006; 70(5): 720-736.

Groves, RM and Peytcheva, E. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis" Public Opinion Quarterly. 2008;72:167–89.

McMillan, S and Culver III, R, "Nonresponse Bias Analysis for Wave 1 of the 2008 Survey of Income and Program Participation (SIPP)", Techincal Report, 2013, available at https://www.census.gov/programs-surveys/sipp/tech-documentation/nonresponse-reports/nonresponse-reports.html.

National Academies of Sciences, Engineering, and Medicine. (2018). *The 2014 Redesign of the Survey of Income and Program Participation: An Assessment.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24864.

Perez-Lopez, D and Welton, A, "Haven't We Met Before? Leveraging Existing Household Survey Data to Evaluate the Effect of Incentives", Poster presented at 2018 AAPOR meetings, Denver, CO. Available at https://www.census.gov/newsroom/press-kits/2018/aapor.html.

Rossmann, J and Gummer, T, "Using Paradata to Predict and Correct for Panel Attrition" *Social Science Computer Review*, 2015.

Singer, E. "Introduction: Nonresponse Bias in Household Surveys." Public Opinion Quarterly 70, no. 5 (2006): 637-645. Available at doi: 10.1093/poq/nfl034.

Sundukchi, M, Westra, A, and Nwaoha-Brown, F, "Source and Accuracy Statement for the 2014 Panel Wave 1 Public Use Files of the Survey of Income and Program Participation (SIPP)", Technical Report, 2017, available at https://census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements/source-accuracy-statements-2014.html.

Survey of Income and Program Participation (SIPP) 2001 Panel Preliminary Wave 1 Core Microdata File, [machine-readable data file] / conducted by the U.S. Bureau of the Census. Washington: The

Bureau [producer and distributor], 2005. Available at
https://www.nber.org/sipp/2001/sipp01t1.pdf

Survey of Income and Program Participation (SIPP) 2004 Panel Preliminary Wave 1 Core Microdata File,
[machine-readable data file] / conducted by the U.S. Bureau of the Census. Washington: The
Bureau [producer and distributor], 2005. Available at
https://www.nber.org/sipp/2004/sipp04w1prel.pdf

Westra, A and Nwaoha-Brown, F, "Nonresponse Bias Analysis for Wave 1 of the Survey of Income and
Program Participation (SIPP)", Technical Report, 2017, available at https://census.gov/programs-
surveys/sipp/tech-documentation/nonresponse-reports/2014nonresponse-reports.html.