ADJUSTING FOR WAVE 1 UNIT NONRESPONSE IN THE 1996 SIPP PANEL

Aref Dajani, Franklin Winters, U.S. Bureau of the Census[1]
Aref N. Dajani, U.S. Bureau of the Census, Room 3089-3, Mail Stop 8700, Washington, DC 20233

Key words: Unit Nonresponse, Attrition, Association Measures, Polytomous Logistic Regression

Introduction and Background:

For the first time since the Survey of Income and Program Participation was launched in 1984, the revised 1996 Redesign focuses primarily on longitudinal measures in a four-year panel with interviews conducted three times a year. As the level of panel-length unit nonresponse and attrition with this design will be expected to be higher than with the previous design, understandable concerns have been raised about nonresponse bias. The level of unit nonresponse and attrition has also grown across waves in previous SIPP panels.[2]

The focus of this paper is SIPP Wave 1 nonresponse. SIPP Wave 1 nonrespondents are not contacted in subsequent waves and the Wave 1 nonresponse adjustment is an integral part of each future wave's weighting adjustment. Because of this, the quality of Wave 1 nonresponse adjustment is a high priority, especially in light of a four-year panel.

We know that the current noninterview categories do not fully account for attrition of low income households. This is based on conclusions of previous studies associated with the 1984 and 1985 Panels that were conducted to evaluate the effectiveness of the SIPP nonresponse adjustment procedure in reducing bias. We also know from these studies that inclusion of the current nonresponse adjustment, relative to no adjustment, results in the reduction of nonresponse bias for estimates of monthly mean and median income at both the household level and the person level.[3]

Despite our weighting procedures, research results show that we are still presently unsuccessful at reducing nonresponse bias for the previous two-year overlapping panels for poverty estimates, marriage estimates, health insurance coverage, migration, and others.

As attrition biases can be examined using existing SIPP data by comparing information from earlier waves to later waves in the same panel, external information needs to be considered when analyzing Wave 1 nonresponse as no prior information in SIPP is available for these respondents.[4] Therefore, the major objectives of this research effort are as follows:

1.   to reduce nonresponse bias by collecting information about nonrespondents for use in Wave 1 nonresponse adjustment, and

2.   to obtain other information about SIPP households to help in assessing nonresponse bias at Wave 1, where very little is known.

This paper documents the first steps in this research effort and presents initial findings. We will discuss the survey instruments used in this study, document how well responses are associated across the instruments, and analyze the relationship of the respondents in this study to the respondents in the 1996 Wave 1 production database. We do this by comparing the distributions of demographic and household characteristics between the respondents in the 1996 Wave 1 production database and the nonrespondents in the 1996 Wave 1 production database who participated in this study. In this way, we will determine whether there are differences between respondents and nonrespondents that can be incorporated into future nonresponse adjustment.

This paper consists of four sections. Section I will document the study design, highlighting the unique sources of data acquired to conduct this study. Section II will describe the methodology for measuring association between various estimators and for testing for equality in distributions for those estimators. Section III will provide the analysis and results and Section IV will offer conclusions, recommendations, and directions for future research.

I. Study Design

To assess the feasibility of improving our adjustment for Wave 1 unit nonresponse in the 1996 Panel, we conducted two surveys. One survey was filled in by field representatives (FRs) after each noninterview to collect observational information; the other was a mail-out/mail-back questionnaire to gather limited information from nonrespondents. Together, the surveys provide insight into the quality of the Wave 1 nonresponse adjustment that is expected to provide information to reduce bias in the estimates highlighted above.

There are many causes of nonresponse in any survey. During Wave 1, a type of nonresponse collectively

categorized in SIPP as Type A nonresponse is targeted. Examples of Type A nonresponses are finding no one home, facing a household refusal, or uncovering a language problem.

Our goal in this study was to obtain response rates that were as high as possible and to keep the incremental workload for the FRs as low as possible. In that effort, we kept the questionnaires short (one page front and back for both FRs and nonrespondents) and changed the mode of the survey instrument from personal interview to mail-out/mail-back. We limited the questions to those items used for Wave 2+ nonresponse adjustment and specific measures of interest concerning income, poverty, and program participation. The following are examples of specific questions:

o   Last month, did anyone in this household work for pay as self-employed?

o   Last month, did anyone in this household receive income from foster child payments?

o   Last month, did anyone receive pension income from Social Security?

o   Last month, did anyone get income from interest from bonds or bank accounts?

o   Does anyone in your household have health insurance with Medicaid?

We also asked questions of the nonrespondents that we believed the FRs could also answer. The following are examples of specific questions that we asked both nonrespondents and FRs:

o   Race of reference person
        (White; Black; American Indian, Eskimo, or Aleut; Asian or Pacific Islander)

o   Monthly household income

o   Rented in public housing project? (yes, no)

o   Number of adults age 15 or older in household

For monthly household income, three categories (<$1200, $1200-$3999, and $4000+) were provided to FRs while six categories (<$500, $500-$1199, $1200-$2999, $3000-$3999, $4000-$8999, and $9000+) were provided to nonrespondents. This was done under the assumption that nonrespondents could provide a more precise estimate of their monthly household income than FRs. The six categories were later collapsed into the three FR categories

when comparing answers between nonrespondents and FRs or when consolidating responses to compare against the 1996 Wave 1 production database.

Nonrespondents

The study of Wave 1 nonrespondents in the 1996 Panel included follow-up of all Wave 1 nonrespondents. All nonrespondents were included because characteristics by subgroup were needed to define more detailed weighting adjustments. By including all nonrespondents we could better analyze bias of subgroup characteristics. When we sent out the nonrespondent questionnaires initially, we did not know how well they would respond to the questionnaires as they either refused a personal interview or were never available for an interview previously. Questionnaires were sent to the Data Processing facility in Jeffersonville, Indiana and a toll-free telephone line was established there to handle questions.

Field Representatives

Field representatives also completed questionnaires for all Wave 1 nonrespondent households. They were instructed to talk to neighbors whenever possible as we wanted as much information as possible gathered from both nonrespondents and FRs for each household. This would enable an evaluation of how well FR responses replicated those of nonrespondents who responded to the questionnaire on those items that were intentionally duplicated across questionnaires. Questionnaires were sent from the FRs to their respective field offices and later forwarded to Census Headquarters. Though the data collection effort did not specifically capture the source or proxy used to fill out each questionnaire, we instructed the FRs to utilize information from the best sources available to them.

II. Methodology

To evaluate whether the results of this study can be used in future research to reduce the nonresponse bias associated with important subject matter estimates such as poverty, program participation, and health insurance coverage, we conducted the following analysis:

We first compared responses across the two surveys to determine how well FRs served as a resource in imputing for nonresponse. High correlations indicate that FRs serve well as such a resource. We calculated three measures of association: a nonparametric percent concordance, a continuous simple correlation, and a categorical Cramer V association measure.[5]

We then consolidated the two surveys, taking answers from the nonrespondent when we had them, taking answers from the FR when we did not, and analyzed distributional properties of respondents and nonrespondents. This was performed at the aggregate level. We produced crosstabulations of key characteristics by their nonrespondent status. FRs filled out questionnaires for Type A nonrespondents who, after further follow-up, may have been converted to a completed response. Because of this, we were able to partition respondents and nonrespondents into three categories:

o   "Early Respondents" are those households that responded in the 1996 Wave 1 production database and did not have FR questionnaires filled out.

o   "Late Respondents" are those households that responded in the 1996 Wave 1 production database and had FR questionnaires filled out because they were originally Type A noninterviews.

o   "Type A Noninterviews" are those households that have completed FR or nonrespondent questionnaires and are in the 1996 Wave 1 production database as Type A noninterviews.

Distributions of demographic and housing factors such as tenure, race, and income were compared for the three types of respondents using a polytomous logistic regression, discussed in [2] and [3]. The higher the log odds ratio in absolute value, the stronger the relationship between the demographic or housing factor and whether the household was an early respondent, a late respondent, or a Type A noninterview.

The objective here was to determine whether it was possible to correct or weight for nonresponse bias by examining the characteristics of respondents who are reluctant to participate in the initial phases of SIPP but later consent and characteristics of individuals who remain nonrespondents. This analysis should help to identify other variables to be used in developing a new nonresponse adjustment procedure, where original SIPP sample and respondents to the study differ.

III. Analysis and Results

Unit Response Rates

The 76% response rate of the FRs was lower than expected as an FR questionnaire was expected for every Type A nonrespondent. On the other hand, the 40% response rate

of the nonrespondents was much higher than expected as these were people who did not respond in the past. Out of 4,579 questionnaires received at Census Headquarters, 2,435 field questionnaires were sent back with identifier data that was matchable to the 1996 Wave 1 production database as Type A noninterviews and 1,290 nonrespondent questionnaires were sent back with identifier data that were matchable to the 1996 Wave 1 production database as Type A noninterviews.

Among the nonrespondent questionnaires, 20% were undeliverable as addressed and 3% were out of scope where the respondents reported not living at the household the previous month. Leaving aside the 21% that were returned to us blank, the remaining 55% of the returned questionnaires had an outcome code of complete or partially complete. We classified questionnaires as complete if at least half of the questionnaire items were filled in and we classified questionnaires as partially complete if any of the questionnaire items were filled in. These 716 questionnaires may only comprise 22% of the 3,194 nonresponses to the original survey, but these were in fact nonrespondents: households that refused to talk with us previously or were otherwise unavailable for interview.

Upon matching the nonrespondent to the field questionnaires, we found that the FRs did a good job in supplementing for remaining nonresponse. While 29% of the nonrespondent questionnaires had an outcome code of complete, 94% of the FR questionnaires had an outcome code of complete. The vast majority of nonrespondent questionnaires were accompanied by matching FR questionnaires.

Item Response Rates

For those 2,435 FR questionnaires sent back with good identifier data, the item response rate was at least 80% for all items. So if the FR sent back a questionnaire, it was most likely filled in. For those who responded, the nonrespondents did not fare too badly, with item response rates in the low 50% range for virtually all items. They even responded well to our income category question with an item response rate of 47%.

Measures of Association Across the Two Surveys

In comparing estimates across the two surveys, the display below shows the results of correlation analyses of eight items in common. These measures of association are based on the households for which both questionnaires were completed. The number of pairs of questionnaires used in calculating each measure of association is also displayed below. All hypotheses involving the correlation

coefficient testing for no association were rejected with p-values of 0.0001. This statistical significance is not surprising as the sample sizes are large.

[The public housing and rent subsidy questions were only asked of renters, resulting in the relatively low number of questionnaires.]

It is difficult to generalize these results to all FR and nonrespondent questionnaires as not all FR questionnaires were accompanied by nonrespondent questionnaires and not all nonrespondent questionnaires were accompanied by FR questionnaires. We consider this an important caveat as we use these measures of association as a basis for inferring whether FR responses are useful when nonrespondent questionnaires are not available.

As mentioned earlier, all three are association measures: percent concordance is a nonparametric measure, simple correlation is a continuous measure, and Cramer V is a categorical measure. As can be seen below, the simple correlation and Cramer V point to the same items as relatively higher or lower than the others. The question is whether those measures show a stronger association relative to the percent concordance measure. We will now go into detail about findings concerning race, income, and public housing.

Race: The high percent concordance occurred as most reference persons in both surveys were categorized as White. The percent correlation and Cramer V statistics were also relatively high because when FRs and nonrespondents classified reference persons as Nonwhite, they agreed on the specific race (Black; American Indian, Eskimo, or Aleut; Asian and Pacific Islander) for virtually every case.

Monthly Household Income: All three measures were relatively low because there were many matched households where the response of the nonrespondent was not the same as that of the FR. Specifically, most nonrespondent low and high income households were classified by the FRs as medium income.

Rented in Public Housing Project?: The high percent concordance occurred as most households were not reported as being rented in public housing projects, which is understandable. The relatively low correlation and Cramer V statistics occurred because few households had both questionnaires placing them in public housing projects. If one questionnaire classified a household in a housing project, it was more likely that the other questionnaire did not.

Comparing Distributions of Demographic and Housing Factors

Having matched FR and nonrespondent questionnaires to each other, distributions of demographic and housing factors for the three categories of respondents and nonrespondents can be compared:

o   Type A Noninterviews -- nonrespondents and FR questionnaires that matched to Type A noninterviewed households in the 1996 Wave 1 production database.

o   Late Responses -- nonrespondents and FR questionnaires that matched to completed interview households in the 1996 Wave 1 production database.

o   Early Responses -- all other completed interview households in the 1996 Wave 1 production database

We now consider the margins, the univariate distribution of demographic and housing factors separately for Type A noninterviews, late responses, and early responses. To generate distributions for late and early responses, we use responses from the 1996 Wave 1 production database as we believe those responses to be more accurate than those of the FRs. To generate distributions for Type A noninterviews, we use the responses for the nonrespondents when we have them and the FRs when we do not as we expect those responses to be more accurate than those of the 1996 Wave 1 production database. Then,

Item Correlation Analysis	Number of Questionnaires	Percent Concordance	Simple Correlation	Cramer V
Number of adults in household	425	72%	0.613	0.508
Number of children in household	408	77%	0.622	0.485
Number of residents in household	425	69%	0.658	0.568
Race of reference person	419	92%	0.677	0.676
Tenure (owner, renter, occupied)	413	84%	0.568	0.484
Rented in public housing project?	194	92%	0.295	0.295
Received rent subsidy?	145	92%	0.355	0.355
Monthly household income	349	52%	0.333	0.277

by using a polytomous logistic regression, we test whether the distributions of the demographic and housing factors are identical. In the logistic regression, we calculate a "-2 Log L" (minus two multiplied by the log of the likelihood ratio) chi-square test statistic to test whether the distributions are identical. Information on polytomous logistic regression is well documented in [2] and [3]. All hypotheses of equality of distribution are rejected with p-values of 0.0001. Again, this statistical significance is not surprising as the sample sizes are large.

Since all factors are statistically significant, we determined which levels of each significant factor contributed to the differences between groups. This was done by looking at the log odds ratios ("b" parameters in the polytomous logistic regression model) to see which ones were largest in absolute value to determine which factors and levels were practically significant. The display below summarizes which factors and levels have the largest log odds in absolute value.

These results imply that nonrespondents tend to be single person, female renters who do not live in public housing projects. It is already known that the number of residents and tenure are associated with response as both factors are currently used in adjusting for nonresponse in SIPP. Public housing project status and household type are not currently used in adjusting for nonresponse in Wave 1 of SIPP.

IV. Conclusions and Recommendations

Shortening the length and changing the mode of the questionnaire was highly successful. Of 3,194 Type A noninterviews in Wave 1 of the 1996 SIPP Panel, all were sent questionnaires. Counting only those forms that were completely filled in or partially filled by the respondent, we obtained a response of 716 questionnaires or a rate of 22%. Of the remaining nonrespondent questionnaires returned to the Census Bureau, approximately half were undeliverable as addressed or out of scope, so it may be possible to obtain a higher response in the future. The mail-out/mail-back short questionnaire worked well when personal interviews were ineffective.

There are certain variables that lead us to conclude that FRs do as well as nonrespondents in providing information. Those are the variables with relatively high correlation or Cramer V statistics. Those are the following: race of reference person, number of residents in household, and number of children in household. There are three variables with low correlations, leading to the conclusion that FRs do not perform as well as nonrespondents in providing information. Those are the following: rented in public housing project, received rent subsidy, and household monthly income. For other variables, further research is prudent.

This research has lead to interesting observations. First, we did not expect to see the results concerning income when comparing FRs and nonrespondents, since we expected many nonrespondents especially to understate their income. When asked to obtain proxy information or estimate household income themselves, FRs tend to understate household income as well. Second, we did not expect to see the disparity concerning public housing. That leads us to believe that nonrespondents and FRs may have differing definitions of public housing projects. In fact, the FR may even be more correct in their definition if they talked to a knowledgeable respondent like a superintendent. We wish to look into a greater understanding of this issue in the future as it potentially indicates a self-identification problem on the part of respondents and an identification problem on the part of FRs, both of which can be problematic to researchers when conducting any poverty survey.

Extensions for Future Research

Now that we have determined that the number of residents, tenure, household type, and public housing project status are most associated with nonresponse, we believe that a valuable next step is to develop a new nonresponse adjustment procedure using these variables in addition to those already used in nonresponse adjustment. We will construct a test database of all respondents and nonrespondents to Wave 1 and determine which adjustment cells nonrespondent households will reside. We will reweight nonrespondent households in those adjustment cells and compare those weights to their original nonresponse adjustment. We will test whether the

Factor (Level)	Type A Noninterviews	Late Responses	Early Responses	Log Odds ("b")
Number of adults in household (1)	38.4%	32.2%	30.5%	-0.709
Number of residents in household (1)	33.3%	29.4%	26.0%	-0.882
Household type (female-headed)	11.2%	0.3%	0.3%	-3.772
Tenure (rent)	34.7%	31.8%	31.8%	-0.601
Rented in public housing project? (yes)	3.6%	6.5%	8.3%	0.794

values of income, poverty, and program participation estimates show a statistically significant change as a result of the reweighting. We will have the option to weight nonrespondents by late respondents instead of all respondents to account for Type A nonresponse as late respondents may be closer to the nonrespondents in distributions of key statistics.

We will refer to the March supplement of the Current Population Survey (CPS) as it is currently used to calculate the national measures of poverty. We acknowledge that comparing SIPP to CPS arrives at concordance, not unbiasedness. Future research will resolve the effect of reweighting on bias reduction.

As valuable as these questionnaires are in conducting empirical research, the questionnaires also form a basis for qualitative, anecdotal research. Many of the nonrespondents felt no inhibition from writing their own comments on the questionnaires. These comments could possibly give some insight into why certain people continue to refuse to be interviewed. It is also possible to extend this study in the future to incorporate administrative records for nonresponse adjustment. These efforts may be undertaken to validate reporting error. For interviewed cases, we can compare values reported in the survey to values derived from administrative records. We can use auxiliary information where possible (e.g., reinterviews) to determine which measures are biased if they disagree. We can then consider whether the differences are systematic; e.g., due to conceptual or time period differences, and whether such differences could be modeled and form the basis for adjusted values.

Whatever the case, having two questionnaires supplement for nonresponse brings us a long way in understanding who nonrespondents are and how to adjust for them as necessary.

Acknowledgments

The authors thank Leroy Bailey, Dennis Schwanz, Jan Tin, and Patrick Flanagan for their substantial and insightful comments upon reviewing this paper. The authors also thank Grace Lee and Agnes Colbert for their data entry and editing support that made this paper possible.

References

[1]   Kendall, M. and Stuart, A. (1979), *The Advanced Theory of Statistics, Volume 2*, New York: Macmillan Publishing Company, Inc., pg. 588.

[2]   DeMaris, Alfred (1995), "A Tutorial in Logistic Regression", Journal of Marriage and the Family, 57, pp. 956-968.

[3]   Hosmer, D.W., Jr. and Lemeshow, S. (1989), *Applied Logistic Regression*, New York: John Wiley & Sons, Inc., pp. 216-238.

[4]   King, Karen E. (1995), "SIPP 1986 - 1993: Monitoring the Effects of Telephone Interviewing - An Analysis of Type A and Type D Noninterview Rates," Census Bureau Memorandum dated February 16, 1995.

[5]   Sanchez, Melissa A. (1991), "SIPP 85: Evaluation of the Nonresponse Adjustment Procedure," Census Bureau Memorandum dated June 24, 1991.

[6]   Tin, Jan (1995), "A Model of Attrition and Income for Dynamic Longitudinal Surveys," Applied Economics, 27, pp. 705-717.

[7]   Tin, Jan (1996), "Program Participation and Attrition: the Empirical Evidence", American Statistical Association 1996 Proceedings of the Section on Survey Research Methods, pp. 669-674.

[8]   Zabel, Jeffrey (1993) "An Analysis of Attrition in the PSID and SIPP with an Application to a Model of Labor Market Behavior," SIPP Working Paper Series No. 9403, Washington, DC: U.S. Bureau of the Census.

[1]   This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributed to the authors and do not reflect those of the Census Bureau.

[2]   See [4] for further discussion on this issue.

[3]   Details of these studies are discussed in [5].

[4]   Issues of attrition bias are addressed in [6], [7], and [8].

[5]   The Cramer V association measure is described in [1] and the percent concordance measure is one minus the gross difference rate, commonly used in reinterview analysis.