

**THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION**

**OVERVIEW OF SIPP
NONRESPONSE RESEARCH
DATA**

No. 201

**S. Mack
R. Petroni
Census Bureau**

U. S. Department of Commerce BUREAU OF THE CENSUS

Overview of SIPP Nonresponse Research¹

Stephen Mack and Rita Petroni
U.S. Bureau of the Census
Washington D.C. 20233

For presentation at the Fifth International Workshop on Household
Survey Non-Response, September 26-28, 1994
Ottawa, Ontario

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

I. Introduction

Nonresponse is an important source of bias for longitudinal surveys. Longitudinal surveys like the Survey of Income and Program Participation (SIPP) require a number of interviews over a period of years. The rate of household nonresponse increases with each successive interview. The characteristics of nonrespondents are likely to differ from respondents. Consequently, the sample becomes less representative of the population over time.

The SIPP gathers information about the financial circumstances of persons, families, and households in the noninstitutionalized U.S. population. Survey participants are asked about cash and noncash income, assets, participation in government assistance programs, employment status, and other items related to their economic situation.

The sample is divided into four rotation groups of equal size. Field Representatives (FRs) attempt to obtain interviews from sample households from one rotation group each month. Each interview of all four rotation groups is called a wave. Starting with the 1984 panel, a panel has typically been introduced each year through 1993. The number of sample households and waves vary with each panel. The 1987 panel, for example, was interviewed for seven waves and started with 12,500 interviews of households in wave 1.

Starting with the 1996 panel, we plan to introduce panels of approximately 50,000 households every four years. The size of 1996 panel will likely be cut to around 45,000 households for budgetary reasons. Sample households will be interviewed every 4 months for about 4 years.

Researchers (Short and McArthur 1986; Petroni 1987; and others) have documented the potential for attrition bias through differential attrition. Other researchers have found differences between estimates of some items from SIPP data and estimates from other sources. For example, marriage rates (Hernandez, 1989) and migration (DeAre 1990).

From the beginning of the SIPP, the Census Bureau has conducted research to measure bias, reduce nonresponse, and better compensate for nonresponse in weighting. Results from three such efforts: Allen and Petroni (1994), Folsom and Witt (1994), and Rizzo et al. (1994); each focusing on alternative weighting adjustments for nonresponse, have recently been reported. This paper provides a short summary of these studies.

II. Overview of Current Weighting

The SIPP provides weights for persons that are appropriate for different types of analysis. Cross-sectional weights are provided for analysis of data from a particular month. Longitudinal weights are provided for analysis of persons over a particular calendar year or over the life of the

panel. All persons classified as interviewed for the appropriate period receive a positive weight. Interviewed persons are persons who were self or proxy respondents for each month in the weighting period for which they were eligible to be interviewed. Persons who die or move to ineligible addresses are no longer eligible to be interviewed. SIPP weights are the product of the following components:

Cross-sectional Weights

- Base Weight (BW) - The inverted probability of selection of a person's household
- Duplication Control Factor (DCF) - Adjusts for subsampling done in the field when the number of sample units is much larger than expected
- Household Noninterview Adjustment Factor (F_N) - Adjusts for noninterviewed households that were eligible for interviews
- First Stage Adjustment Factor (F_{1S}) - To reduce between PSU variance, this factor was found to be ineffective and will not be used for 1996+ panels
- Second Stage Adjustment Factor (F_{2S}) - To adjust estimates to population controls and cause husbands and wives weights to be equal

The final cross-sectional weight is

$$FW_C = BW \times DCF \times F_N \times F_{1S} \times F_{2S}$$

Longitudinal Weights (panel and calendar year)

- Initial Weight (IW) - Cross-sectional weight before second stage adjustment from the appropriate panel or calendar year "control" month
- Noninterview Adjustment Factor (F_L) - Adjusts for noninterviewed persons with nonzero initial weights
- Second Stage Adjustment Factor (F_{2S}) - To adjust estimates to population controls. Husbands and wives weights are not equalized.

The final longitudinal weight is

$$FW_L = IW \times F_L \times F_{2S}$$

Additional factors were used in some panels to adjust for special situations. For example, a "Sample Cut Adjustment Factor" was used to adjust longitudinal weights of the 1985 panel for the February 1986 sample cut.

III. Mover Nonresponse Adjustment Research (longitudinal weighting)

Allen and Petroni (1994) conducted research on incorporating two alternative mover's adjustments into the SIPP weighting. The research stems from earlier work by Petroni (1993) which was undertaken to address concerns about levels of SIPP nonresponse and bias resulting from nonresponse (Hernandez 1989,1990; Hill 1993,1994; etc.).

For this research, they used the 1987 SIPP panel data and modified the longitudinal weighting procedures to produce panel and calendar year weights. For the first alternative, the researchers computed final weights as $IW \times F'_L \times F_{2S}$. They computed the current nonresponse adjustment procedure separately for movers and nonmovers to determine F'_L .

For the second alternative, the researchers computed final weights as $IW \times F_L \times F'_{2S}$. F'_{2S} was the result of 30 iterations. Each iteration consisted of first ratio adjusting to CPS estimates of movers and nonmovers by age, race, and sex and then calculating

Compared with SIPP estimates based on original SIPP weights, the alternative weights produced mover estimates which were closer to benchmarks estimates. However, alternative 1 estimates were always statistically different from the benchmarks and alternative 2 estimates sometimes were and sometimes were not.

For other estimates - number of marriages, number of divorces, percent below the poverty level, and median family income - both alternatives produced estimates which were numerically, yet not always statistically, closer to the benchmark estimates than were estimates from original SIPP weights. For most estimates, all three weights produce estimates that are either all statistically different or all not statistically different from the benchmark estimate.

Benchmark estimates such as CPS income from certain sources may be more biased than SIPP estimates due to differences in methodology (Jabine, et al., 1990). For example, the SIPP recall length is four months. The CPS recall length is twelve months. Other methodological differences may also create differences between benchmark and SIPP estimates.

Because the results provide no strong evidence that either alternative reduces biases, we do not plan to incorporate either alternative into weighting procedures and no further research is planned.

IV. Constrained Response Propensity Adjustment for Panel Nonresponse

A new report from the Research Triangle Institute (RTI), Folsom and Witt (1994) proposes a longitudinal nonresponse weighting adjustment based on response propensity models. The RTI model was fitted to 1987 panel data and used to adjust the initial panel weights for wave 2+ nonresponse.

The authors modeled the response propensity of persons who responded to the initial interview of the 1987 panel. Constrained logistic regression models were fitted to each of seven nonresponse classes. The nonresponse classes were defined in terms of average household income, race/ethnicity, marital status, and census region. The response propensity models had the form

$$\hat{\lambda}_i = 1 + \exp\left(-\sum_j x_{ij} \hat{\beta}_j\right)$$

subject to the generalized raking constraints:

$$\sum_i IW_i \lambda_i r_i x_{ij} = \sum_i IW_i x_{ij}$$

where x_{ij} = the wave one value of covariate j for person i
 λ_i = the nonresponse adjustment factor
and r_i = the zero-one longitudinal response indicator

Another constraint, $\hat{\lambda}_i \leq 2$, was incorporated to control variation in the adjusted weights,

$$IW_i \hat{\lambda}_i$$

, to an acceptable level. Note that wave one respondents are considered longitudinal respondents only if they responded to all subsequent interviews for which they were eligible.

The final RTI adjusted weights were formed by performing the standard second stage adjustment of the weights population controls. The final RTI adjusted weight was

$$FW_{RTI}(i) = IW_i \hat{\lambda}_i F_{2s}(i), \quad \text{for each respondent } i$$

Folsom and Witt compared various statistics using RTI revised weights and original 1987 panel weights with the corresponding 1989 panel wave 1 estimates. Table 1 shows some of these comparisons. The expectation was that if the RTI revised weights had reduced the bias from

wave 2+ nonresponse, then the RTI revised estimates would be closer to the 1989 wave 1 estimates than the corresponding estimates made with the original 1987 panel weights.

Table 1. Comparison of January, 1989 Estimates Using 1987 Panel Original Longitudinal Weights, 1987 Panel RTI Revised Longitudinal Weights, and 1989 Panel Wave 1 Weights.

Characteristic	Estimate			% Relative Difference from Benchmark	
	Original	RTI Revised	Benchmark (1989 Panel)	Original	RTI Revised
Total Population					
Mean Nonzero \$ Amounts					
Personal Income	1455.97 ^a	1460.71	1447.94	0.55	0.88
Family Income	2983.70	2981.95 ^b	2903.15 ^c	2.77	2.71 *
Unemployment	519.04	521.91 ^b	461.68 ^c	12.42	13.05
Food Stamps	130.21 ^a	132.92	128.08	1.66	3.78
AFDC	353.20 ^a	357.38	362.76	-2.63	-1.48 *
Proportion with Nonzero Amounts					
Personal Income	73.34 ^a	73.18	72.94	0.55	0.34 *
Family Income	98.18 ^a	98.13 ^b	98.86 ^c	-0.69	-0.75
Unemployment	1.00	1.01	0.94	6.12	7.29
Food Stamps	2.62 ^a	2.52	2.59	1.29	-2.70
AFDC	1.13	1.12 ^b	1.33 ^c	-15.23	-15.57
Poverty Ratios					
[0.00,1.00)	13.13	13.20	13.68	-4.02	-3.52 *
[1.00,1.25)	4.65	4.65	4.55	2.11	1.98 *
[1.25,1.50)	4.97	4.98	4.95	0.23	0.48
[1.50,2.00)	10.23	10.25	10.90	-6.17	-6.03 *
[2.00,3.00)	20.20	20.19	20.56	-1.75	-1.82
3.00 or greater	46.82	46.75 ^b	45.35 ^c	3.26	3.09 *
Coverage					
Food Stamps	6.84 ^a	6.69	6.28	8.88	6.51 *
AFDC	3.17	3.15	3.55	-10.95	-11.33
In Labor Force (age 16+)	66.18 ^a	65.95	66.21	-0.05	-0.38
Bonds	20.23 ^a	20.36 ^b	32.96 ^c	-38.62	-38.22 *

* Relative difference of RTI estimate is closer to zero than relative difference of original estimate.

^a Original estimate is significantly different from RTI revised estimate at 10% level or less.

^b RTI revised estimate is significantly different from benchmark estimate at 10% level or less.

^c Original estimate is significantly different from benchmark estimate at 10% level or less.

Unfortunately, the evaluation did not clearly demonstrate any reduction of nonresponse bias from the response propensity model approach.

There are no current plans to do any further SIPP research on inverse nonresponse propensity weighting adjustment. The response propensity / generalized raking methodology does, however, have some appealing properties:

- Variance inflation can be controlled by bounds on weight adjustments
- Larger numbers of predictor variables can be used with inverse response propensity adjustments than with the standard SIPP nonresponse methodology.

V. A Comparison of Alternative Forms of Panel Nonresponse Adjustment

A report by Rizzo, Kalton, Brick, and Petroni (1994) on Westat research compares several alternative panel nonresponse adjustments using data from the 1987 panel. Nonresponse adjustments based on response propensity models, generalized raking, and observed response rates were compared with the original SIPP panel weights. Two of the adjustments used the CHAID algorithm to define nonresponse adjustment cells. The CHAID algorithm sequentially splits the data into subgroups according to whatever variable has the largest response rate differences between its categories. Selected estimates of program participation, etcetera, were made using the alternative adjustments and compared. Estimates were also compared to estimates from 1989 panel wave one data and to benchmarks from published administrative data.

Variables from wave one responses that were related to panel nonresponse were first chosen. Ten variables were selected to use in the nonresponse adjustments based on an initial screening and logistic regression models. These variables were used in the raking, logistic regression, logistic regression / observed, and collapsed cell alternatives. An additional three variables were selected by the CHAID algorithm. The additional variables were used only in the CHAID alternatives, CHAID I and CHAID II. The variables are defined in table 2.

Table 2. Variables used in Westat alternative nonresponse adjustments.

Variable	Categories
Selected by Screening and Logistic Regression Analysis	
1. Age	<16,, 16-24, 25-50, 51-71, 72+
2. Race	white, black, other
3. Relationship to Reference Person in Household	primary family member, other
4. Census Region	New England, Mid Atlantic, South Atlantic, East South Central, West/East North Central, Mountain/West South Central
5. Tenure	home owner, renter, other
6. Number of Wave One Items Imputed	0, 1, 2 or 3, 4 or more
7. Bond Status	no bonds, some bonds
8. Layoff Status	laid off during wave one, not laid off
9. Food Stamps	recipient, nonrecipient
10. Class of Work	business, government, other
Selected by CHAID Analysis	
11. Sex	male, female
12. Monthly Household Income	less than \$1,200, \$1,200 to \$8,000, more than \$8,000
13. Education	highest grade completed was tenth or eleventh grade, other

The alternative panel nonresponse adjustments are:

Raking Adjustment: Initial weights were raked to equal the wave one marginal distribution of variables 1 through 10. The marginal totals were the sum of wave one weights for all persons, i.e. panel respondents and panel nonrespondents.

Logistic Regression: Initial weights were adjusted for nonresponse based on the predicted response rate in each cell. The predicted response rate was based on the main effects logistic regression model using variables 1-10.

Logistic Regression / Observed: Initial weights were adjusted for nonresponse based on observed response rates for cells with 25 or more observations and predicted response rates otherwise. The predicted response rate was based on the main effects logistic regression model using variables 1-10.

Collapsed Cells: Initial weights were adjusted for nonresponse based on observed response rates. Cells were defined by variables 1-10. Cells were combined with other cells until each cell had more than 30 observations. Cells with similar predicted response rates were combined.

Predicted response rates were based on the main effects logistic regression model using variables 1-10.

CHAID I: Adjustment cells were defined using the CHAID algorithm and variables 1-7 and 11. Cells were required to have at least 25 observations. The CHAID I model ended with 99 adjustment cells. The nonresponse adjustment was based on the observed response rate within each cell.

CHAID II: Adjustment cells were defined using the CHAID algorithm and variables 1-13. Cells were required to have at least 25 observations. The CHAID II model resulted in 142 adjustment cells. Initial weights were adjusted for nonresponse according to observed response rates within cells.

The six sets of alternative nonresponse adjusted weights were poststratified to CPS population controls. The Westat poststratification procedure differed somewhat from the second stage adjustment used with the original '87 panel weights, but the differences were considered as minor.

Table 3 presents some comparisons of estimates from the 1987 SIPP panel, the six Westat alternative weights, wave 1 of the 1989 SIPP panel, and benchmarks from other sources. The 1989 SIPP data is also considered a benchmark because it is not subject to wave 2+ nonresponse. Note that table 3 contains only a portion of the comparisons from the original study.

The estimates from the original SIPP panel weights and the six alternative nonresponse treatments were all very similar to each other. There were greater differences between the seven estimates from the 1987 panel and the benchmark estimates (which come from other sources). The results showed that none of seven nonresponse adjustments were better than the others at reducing panel nonresponse bias.

Table 3. Comparison of 1987 SIPP Panel Estimates Made Using Alternative Nonresponse Adjustments from Rizzo, et al. (1994).

	Original '87 Panel Weights	Raking Adjustment	Logistic Regression	Logistic Regression / Observed	Collapsed Cells	CHAID I	CHAID II	89 Panel, Wave 1	Benchmark
Total Population - January 1989									
AFDC	3.10	3.10	3.12	3.14	3.12	3.14	3.02	3.56	4.24 ^a
Food Stamps	6.71	6.58	6.63	6.67	6.64	6.70	6.59	6.30	7.29 ^b
SSI	1.65	1.66	1.67	1.66	1.64	1.66	1.61	1.65	1.74 ^a
Poverty Rate	12.91	12.93	12.98	13.02	12.97	12.99	12.91	14.46	
Median household - income	2601	2602	2600	2597	2607	2607	2607	2550	
Population Age 15+, January 1989									
Employed	62.74	62.42	62.36	62.34	62.43	62.42	62.52	61.60	
Unemployed	3.57	3.63	3.64	3.63	3.60	3.58	3.60	4.52	
Not in Labor Force	33.69	33.95	34.01	34.03	33.96	34.01	33.88	33.88	
Total Population 1987									
Marriages	1.39	1.41	1.41	1.40	1.39	1.39	1.39		1.86 ^c
Divorces	0.51	0.49	0.50	0.50	0.49	0.50	0.51		0.90 ^c
Movers	12.88	13.33	13.32	13.32	13.19	13.36	13.37		17.99 ^d

^a Social Security Bulletin, Volume 52, No. 3.

^b USDA Food and Nutrition Service, National Data Bank, unpublished data.

^c National Center for Health Statistics: Vital Statistics of the U.S., 1987, Volume III, Marriage and Divorce, DHHS Pub. No. (PHS) 91-1103.

^d U.S. Bureau of the Census, Current Population Reports, Population Characteristics, P-20, No. 473.

VI. Other SIPP Research

The Census Bureau is continuing to sponsor research aimed at reducing nonresponse bias in the SIPP. New methods to reduce mover nonresponse are being investigated (Allen 1994). Research on additional weighting alternatives to compensate for nonresponse is underway. The Census Bureau is investigating:

- Regression weighting methods as an alternative to the current longitudinal weighting procedure (An, Breidt, and Fuller 1994)
- Using Internal Revenue Service (IRS) data to improve income estimates (Dorinski and Huang 1994). The research focuses on using IRS data in addition to population controls in the second stage adjustment procedure for calendar year weights.

Recent and planned changes to SIPP weighting are being made based on previous research. These changes include:

- We will use some new variables in the nonresponse adjustment for wave 1 and wave 2+ weights starting with the 1996 panel. This revision is partially attributable to a number of studies.
- Single wave interim household nonresponse is now imputed rather than adjusted for in longitudinal weights. This change began with the 1991 panel and is based on research by Lepkowski, Miller, and Luis (1993).
- The first stage weight adjustment will be eliminated starting with the 1996 panel based on research by James (1994).

We will continue to research methods to improve SIPP weights and reduce nonresponse bias.

References

- Allen, T. (1994). *SIPP: Results of Research on Tracking Movers*. Internal Census Bureau Memorandum from Allen to Research and Evaluation Committee, January 31, 1994
- Allen, T., and R. Petroni (1994). Mover Nonresponse Adjustment Research for the Survey of Income and Program Participation. *American Statistical Association 1994 Proceedings of the Section on Survey Research Methods*.
- An, A., F. Breidt, and W. Fuller (1994). Regression Weighting Methods for SIPP Data. *American Statistical Association 1994 Proceedings of the Section on Survey Research Methods*.
- DeAre, D. (1990). Longitudinal Migration Data From the Survey of Income and Program Participation, *Perspectives of Migration Analysis*, Chapter 2 of Current Population Report P-23, #166, U.S. Bureau of the Census
- Dorinski, S., and H. Huang (1994). Use of Administrative Data in SIPP Longitudinal Estimation. *American Statistical Association 1994 Proceedings of the Section on Survey Research Methods*.
- Folsom, R. and M. Witt (1994). Testing a New Attrition Nonresponse Adjustment Method for SIPP. *American Statistical Association 1994 Proceedings of the Section on Survey Research Methods*.
- Hernandez, D. (1989). *Components of Longitudinal Household Change for 1984-85: An Evaluation of National Estimates from the SIPP*. SIPP Working Paper Series, No. 8922, U.S. Bureau of the Census
- Hernandez, D. (1990). *Preliminary Evaluation of Household Change Estimates in 1984, 1985, and 1986 SIPP Panels*. Internal Census Bureau Memorandum from Hernandez to Norton, March 1, 1990
- Hill, D. (1993). *An Investigation of "Under" Estimates in the SIPP*. Survey Research Institute, Toledo, Ohio
- Hill, D. (1994). Weighting for Nonresponse in Event-History Analysis. *American Statistical Association 1994 Proceedings of the Section on Survey Research Methods*.
- Huggins, V., and R. Fay (1988). Use of Administrative Data in SIPP Longitudinal Estimation, *American Statistical Association 1988 Proceedings of the Section on Survey Research Methods*, pp. 354-359

- Jabine, T., K. King, and R. Petroni (1990). *Survey of Income and Program Participation Quality Profile*. U.S. Bureau of the Census
- James, T., (1994). *SIPP 95+: Dropping the First Stage Factor from the SIPP Weights*. Internal Census Bureau Memorandum from Huggins to Cahoon, April 6, 1994
- King, K., (1988a). *SIPP 85+: Cross-Sectional Weighting Specifications for Wave 1*. Internal Census Bureau Memorandum from Waite to Walsh, July 6, 1988
- King, K., (1988b). *SIPP 85+: Cross-Sectional Weighting Specifications for Second and Subsequent Waves -- Revision*. Internal Census Bureau Memorandum from Waite to Walsh, August 4, 1988
- King, K., (1990). *Specifications for Panel File Longitudinal Weighting of Persons*. Internal Census Bureau Memorandum from Waite to Courtland, June 1, 1990
- Lepkowski, J., D. Miller and E. Luis (1993). *Imputation for Wave Nonresponse in the SIPP*. Final Report, Institute for Social Research, University of Michigan, March 1993
- McArthur, E. and K. Short (1986). Life Events and Sample Attrition in the Survey of Income and Program Participation. *American Statistical Association 1986 Proceedings of the Section on Survey Research Methods*, 200-205.
- McArthur, E. (1988). *Measurement of Attrition Through the Completed SIPP 1984 Panel; Preliminary Results*. Internal Census Bureau Memorandum from McArthur to Kasprzyk, March 4, 1988
- Petroni, R. (1987). *SIPP 84: Characteristics of Initially Interviewed Persons by Response Status*. Internal Census Bureau Memorandum from Nonresponse Workgroup for the Record, September 3, 1987
- Petroni, R. (1993). *SIPP Nonresponse Adjustment Research*. Presented at the Fourth International Workshop on Household Survey Nonresponse, September 7-9, 1993, Bath, England.
- Rizzo, L., G. Kalton, J. M. Brick, and R. Petroni (1994). Weighting for Panel Nonresponse in the Survey of Income and Program Participation. *American Statistical Association 1994 Proceedings of the Section on Survey Research Methods*.