# The Effectiveness of Weighting Cell Adjustments for Longitudinal Nonresponse

Leroy Bailey
U.S. Census Bureau

**Abstract**

In this paper longitudinal respondents are considered as members of a panel survey sample for whom data are collected every wave of a given period under consideration, and members of the panel for whom there is nonresponse for at least one wave of the designated period are termed longitudinal nonrespondents. To compensate for the effects of wave nonresponse and attrition, the weights of longitudinal respondents can be adjusted. There is a persistent need to ensure that this weight modification is adaptable to changes in the data structure occasioned by longitudinal nonresponse. The paper will discuss strategies for the selection of longitudinal nonresponse weighting cells, examine the impact of nonresponse weight adjustments on cross-sectional estimates, and provide an evaluation of selected weighting alternatives designed to account for wave nonresponse and attrition. Empirical results from the 1993, 1996, and 2001 panels of the Survey of Income and Program Participation are presented for the selected weighting cell adjustment alternatives.

**Key Words:** attrition; wave nonresponse; longitudinal weighting; logistic regression

## 1. Introduction[1]

For many years survey practice has included the use of weighting cell nonresponse adjustment techniques to account for unit nonresponse. These techniques are widely used as a result of their intuitive appeal and the relative ease with which they can be implemented. Most of the applications of nonresponse weighting cell methods require the partitioning of the sample into homogeneous cells or groups relative to specific population characteristics. Such cells are also referred to as response homogeneity groups or RHGs (Sarndal, Swenson, and Wretman, 1992). In longitudinal surveys the weighting cell methodology can be employed to compensate for the nonresponse encountered in a given wave or round of data collection. Moreover, in addition to providing a rich source of data which permit the study of the dynamic characteristics of a population, longitudinal studies can produce data, not available for cross-sectional surveys, that can be used in the modeling of longitudinal nonresponse and the development of desirable nonresponse bias adjustment methods.

This paper reports the major findings of a study of the effectiveness of alternative weighting cell estimators in accounting for longitudinal nonresponse in the Survey of Income and Program Participation (SIPP). The following section summarizes the theoretical framework for selected weighting schemes to adjust for longitudinal nonresponse. The weighting cell adjustment alternatives considered for the study are presented in Sections 3. Section 4 provides empirical results for cross-sectional estimation, and concluding remarks are offered in Section 5.

## 2. Theoretical Framework for Longitudinal Weighting

The general focus of the study is the effectiveness of selected sample reweighting techniques in accounting for wave nonresponse in estimates of descriptive statistics for a finite population. From the population denoted by $U = \{U_1, U_2, ..., U_N\}$, we select a sample *s* of size n. For survey variable *y*, let $Y_k$ denote the population value for the kth unit of U, and $\pi_k$ be the unit's first order inclusion probability. For the complete response case, assuming no measurement error, we know that an unbiased estimator for the population total $T_y = \Sigma_1^N Y_k$ is $\Sigma_{k \in s} Y_k \pi_k^{-1}$, the Horvitz- Thompson estimator

(HTE). However, in the presence wave nonresponse, the estimator should be modified to account for the missing data.

We assume that for wave t the kth population unit has a nonzero probability $\phi_{kt}$ of responding to the survey, and the unbiased estimator corresponding to

the HTE becomes $\hat{T}_{yt} = \sum_{k \in s_{r(t)}} \frac{Y_{kt}}{\pi_k \phi_{kt}}$, where $s_{r(t)}$ is the set of longitudinal respondents. Since the $\phi_{kt}$ are unknown, we have to resort to modeling to obtain a suitable reweighting estimator designed to reduce the bias occasioned by longitudinal

[1] Disclaimer: This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical, methodological, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

nonresponse. To ensure that the sum of the final survey weights is consistent with known population control totals for specific auxiliary variables, a poststratification ratio adjustment is often made at the final stage of the reweighting process. The resultant expression for the estimator of the total is

$$\hat{T}_{yt} = \sum_{k \in s_{r(t)}} \frac{Y_{kt} g_{kt}(s)}{\pi_k \hat{\phi}_{kt}}, \qquad (2.1)$$

where $g_{kt}(s)$ is an adjustment factor designed to benchmark the weights to population control totals, and $\hat{\phi}_{kt}$ is determined by the estimation methodology applicable to the assumed response model.

## 3. Weighting Cell Adjustment Alternatives

### 3.1 Modeling Response Propensity

The longitudinal nonresponse weighting cells that are defined for survey estimation provide estimates of response probabilities for the sample units they comprise and reflect a response propensity model. For this discussion we will assume that unit level data for various auxiliary variables are available. We define the response model as

$$\phi_k = \phi_k(\mathbf{x},\boldsymbol{\beta}), \qquad (3.1.1)$$

where $\mathbf{x}$ is a p-vector of response predictor variables, and $\boldsymbol{\beta}$ is the p-vector of model parameters.

The conventional nonresponse weighting cell procedures are based on the RHG model for which combinations of response (nonresponse) predictor variables divide the population and the realized sample into mutually exclusive and exhaustive (MEE) categories; the weighting cells correspond to the MEE categories. For the cth weighting cell the estimator of the nonresponse adjustment factor ( under the ignorable nonresponse assumption) for each respondent of the cell is

$$\hat{\phi}_c^{-1} = \tilde{T}_c / \Sigma_{k \in s_{cr}} \pi_k^{-1} = \Sigma_{k \in s_c} \pi_k^{-1} / \Sigma_{k \in s_{cr}} \pi_k^{-1}, \qquad (3.1.2)$$

where $\tilde{T}_c$ is the inflation estimator for the size of the cth weighting cell; $s_c$ is the set of sample cases in the cth weighting cell; and $s_{cr}$ is the corresponding set of sample respondents in the cell.

### 3.2 Selected Alternatives

The alternatives considered for this study are based on logistic regression models, for which we assumed that $\phi_k^{-1} = 1 + \exp(-x_k'\beta)$. For the first alternative, which will subsequently be denoted as LR, a logistic regression model was fitted to the probability of response to the survey, and the inverse of the estimate of the response probability was applied directly to the weights of the longitudinal respondents. For the second alternative, LR/X, the longitudinal nonresponse weighting cells were defined by the covariates of the logistic regression model. Covariates included such variables as education, race, geographic region, homeowner status, and previous-wave value of survey items. The cells were determined by the quintiles of the model for the third alternative (LR/Q). The derived nonresponse weight adjustment for units belonging to a given weighting cell based on the latter two alternatives, like that for the current procedure (CUR), is the inverse of the cell's weighted response rate shown in (3.1.2).

## 4. Empirical Results

### 4.1 Data Sources

The Survey of Income and Program Participation (SIPP) is a longitudinal survey conducted by the U.S. Census Bureau. It produces national-level estimates for the U.S. resident population and subgroups, and has provided the basis for studies and analysis of selected dynamic characteristics.

To compensate for the potential effects of person-level longitudinal nonresponse bias, we adjust the weights of the SIPP longitudinal respondents - members of the survey sample for whom data are collected every wave of the longitudinal period under consideration. This reweighting procedure divides the sample into roughly 150 nonresponse weighting cells, that are assumed homogeneous relative to response propensity and values of the survey variables. The cells are determined prior to the beginning of a SIPP panel, and remain fixed for the duration of the panel. Researchers have investigated the effectiveness of this weight modification. See for example, Rizzo, Kalton, Brick and Petroni (1994), Hendrick (1996) and Slud and Bailey (2006). However, there is a persistent need to ensure that our compensatory procedures are adaptable to changes in the rate of longitudinal nonresponse and that their effectiveness is not significantly diminished.

The primary data sources for the study were the 1993,1996, and 2001 SIPP panels, which included nine, twelve, and nine waves, respectively. The associated sample sizes (households) for the three panels were 21,823; 40,188; and 40,489.

### 4.2 Methodology

In absence of a complete set of reliable benchmarks for the selected survey items, the evaluation of the effectiveness of weighting cell methods in reducing longitudinal nonresponse bias in SIPP entailed comparing Wave 1 estimates, derived from Wave 1 respondents, with Wave 1 estimates based on longitudinal respondents from subsequent waves, whose weights have been adjusted for longitudinal nonresponse. For example, for the 2001 panel we derived separate retrospective estimates for Wave 1 using longitudinally adjusted respondent data from waves 4 and 9. The underlying assumption for the associated comparisons was that the Wave 1-based estimates were "more accurate" than those derived from subsequent waves, and therefore, for the evaluation they could serve as a standard for the estimates from the latter waves.

### 4.3 Results

Tables 4.3 A-C provide comparisons of Wave 1 estimates of selected SIPP totals. Relative differences have been computed for the estimates produced from the longitudinal respondents. The survey items selected for this presentation are AFDC (Aid to Families with Dependent Children), food stamp recipiency, possession of health insurance, and status relative to unemployment and poverty.

Our initial set of comparisons are between the original Wave 1 estimates and the longitudinally adjusted estimates based on the current SIPP weighting cell nonresponse adjustment procedure (CUR). The major conclusion from the examination of these results is that SIPP item estimates may still include significant bias after the longitudinal nonresponse adjustments are made; moreover, in general the effects of longitudinal nonresponse on
SIPP estimates are seemingly exacerbated as time-in-sample is increased. The data shown in these tables are for item totals. The magnitude of the "measures of error" is generally expected to be larger for other domains of interest.

Ostensibly, concurrent with increasing attrition, we have indications of a response mechanism that could potentially result in nonresponse bias not accounted for by the current SIPP longitudinal nonresponse adjustment methodology.

Tables 4.3 A-C also permit a comparison of SIPP estimates based on the current nonresponse weighting cell adjustment methodology (CUR) with estimates from the three other weighting alternatives identified earlier. The general form of the estimators from which the estimates of the tables are derived is given in equation 2.1. Differences in the four alternative estimators essentially result from the different estimators of $\phi_{kt}$, the responsepropensity for the kth unit at Wave t, for the varied response models considered. Equation 2.1 is also the general form of the estimator associated with the Wave 1- based estimates, the standard for the comparisons. However, for the estimator for the population total at Wave 1, denoted by

$$\hat{T}_{y1} = \Sigma_{k \in s_r(1)} Y_{k1} (\pi_k \hat{\phi}_{k1})^{-1} g_{k1}(s), \quad (4.3.1)$$

$\hat{\phi}_{k1}^{-1}$ is a Wave 1 household nonresponse adjustment. For Wave t, t≥2, $\hat{\phi}_{kt}^{-1}$ can be represented as the product of the Wave 1 household nonresponse adjustment and the person-level nonresponse adjustment for the applicable wave.

Table 4.3 A. Comparison of Nonresponse Weighting Alternatives - 1993 Totals
(In Thousands)

| Item | Wave 1 Estimate | Wave 1 Estimation Based on Wave 5 Respondents | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CUR | | LR | | LR/X | | LR/Q | |
| | | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) |
| Food Stamps | 25,812 | 25,821 | 0.03 | 25,767 | -0.17 | 25,882 | 0.27 | 25,229 | -2.26 |
| AFDC | 13,234 | 13,232 | -0.01 | 13,315 | 0.62 | 13,394 | 1.21 | 12,972 | -1.97 |
| Health Insurance | 217,570 | 219,865 | 1.05 | 219,884 | 1.06 | 219,880 | 1.06 | 220,342 | 1.27 |
| In Poverty | 41,119 | 39,909 | -2.94 | 40,832 | -0.70 | 41,065 | -0.13 | 39,459 | -4.04 |
| Unemployed | 9,378 | 9,211 | -1.78 | 8,997 | -4.06 | 9,017 | -3.86 | 8,824 | -5.91 |

| Item | Wave 1 Estimate | Wave 1 Estimation Based on Wave 9 Respondents | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CUR | | LR | | LR/X | | LR/Q | |
| | | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) |
| Food Stamps | 25,812 | 25,700 | -0.43 | 25,857 | 0.18 | 25,874 | 0.24 | 25,985 | 0.67 |
| AFDC | 13,234 | 12,937 | -2.24 | 13,240 | 0.05 | 13,242 | 0.06 | 13,281 | 0.36 |
| Health Insurance | 217,570 | 221,037 | 1.59 | 220,981 | 1.57 | 220,971 | 1.56 | 220,877 | 1.52 |
| In Poverty | 41,119 | 39,874 | -3.03 | 41,306 | 0.46 | 41,393 | 0.67 | 41,368 | 0.61 |
| Unemployed | 9,378 | 9,132 | -2.62 | 8,806 | -6.10 | 8,846 | -5.68 | 8,815 | -6.00 |

Table 4.3 B. Comparison of Nonresponse Weighting Alternatives - 1996 Totals
(In Thousands)

| Item | Wave 1 Estimate | Wave 1 Estimation Based on Wave 4 Respondents | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CUR | | LR | | LR/X | | LR/Q | |
| | | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) |
| Food Stamps | 27,423 | 27,370 | -0.19 | 27,678 | 0.93 | 27,634 | 0.77 | 27,539 | 0.42 |
| AFDC | 14,101 | 14,101 | 0.00 | 14,271 | 1.20 | 14,247 | 1.03 | 14,181 | 0.57 |
| Health Insurance | 194,591 | 195,918 | 0.68 | 195,726 | 0.58 | 195,790 | 0.62 | 195,963 | 0.70 |
| In Poverty | 41,796 | 41,096 | -1.68 | 42,188 | 0.94 | 42,067 | 0.65 | 41,764 | -0.08 |
| Unemployed | 6,406 | 6,182 | -3.49 | 6,588 | 2.85 | 6,539 | 2.08 | 6,294 | -1.75 |

| Item | Wave 1 Estimate | Wave 1 Estimation Based on Wave 12 Respondents | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CUR | | LR | | LR/X | | LR/Q | |
| | | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) |
| Food Stamps | 27,403 | 26,993 | -1.57 | 27,547 | 0.45 | 27,387 | -0.13 | 27,184 | -0.87 |
| AFDC | 14,101 | 13,214 | -5.86 | 13,662 | -3.40 | 13,528 | -4.06 | 12,862 | -3.36 |
| Health Insurance | 194,591 | 197,299 | 1.39 | 196,816 | 1.14 | 197,031 | 1.25 | 197,385 | 1.44 |
| In Poverty | 41,796 | 41,141 | -1.57 | 42,324 | 1.26 | 42,404 | 1.45 | 41,133 | -1.59 |
| Unemployed | 6,406 | 6,046 | -5.62 | 6,948 | 8.47 | 6,386 | -0.30 | 5,934 | -7.36 |

Table 4.3C. Comparison of Nonresponse Weighting Alternatives - 2001 Totals
(In Thousands)

| Item | Wave 1 Estimate | Wave 1 Estimation Based on Wave 4 Respondents | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CUR | | LR | | LR/X | | LR/Q | |
| | | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) |
| Food Stamps | 18,028 | 18,461 | 2.40 | 18,086 | 0.33 | 18,295 | 1.48 | 18,232 | 1.13 |
| AFDC | 5,600 | 5,762 | 2.90 | 5,713 | 2.02 | 5,756 | 2.80 | 5,738 | 2.47 |
| Health Insurance | 211,341 | 212,217 | 0.41 | 212,329 | 0.47 | 212,287 | 0.45 | 212,344 | 0.47 |
| In Poverty | 39,680 | 39,247 | -1.09 | 38,959 | -1.82 | 39,036 | -1.62 | 38,972 | -1.78 |
| Unemployed | 5,562 | 5,442 | -2.15 | 5,379 | -3.29 | 5,385 | -3.18 | 5,378 | -3.09 |

| Item | Wave 1 Estimate | Wave 1 Estimation Based on Wave 9 Respondents | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CUR | | LR | | LR/X | | LR/Q | |
| | | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) | Estimate | Relative Diff. (%) |
| Food Stamps | 18,028 | 18,503 | 2.63 | 18,090 | 0.34 | 18,303 | 1.53 | 18,217 | 1.05 |
| AFDC | 5,600 | 5,807 | 3.70 | 5,753 | 2.73 | 5,797 | 3.52 | 5,772 | 3.07 |
| Health Insurance | 211,341 | 212,273 | 0.44 | 212,415 | 0.51 | 212,357 | 0.48 | 212,438 | 0.52 |
| In Poverty | 39,680 | 39,231 | -1.13 | 38,982 | -1.76 | 39,056 | -1.57 | 38,971 | -1.79 |
| Unemployed | 5,562 | 5,389 | -3.11 | 5,326 | -4.24 | 5,332 | -4.13 | 5,333 | -4.11 |

Results varied by panel, wave, and survey item. In many cases the estimates derived from the current weighting cell procedure and the corresponding estimates based on the alternative methods are not statistically different from each other nor form the Wave 1 standard. However, the table entries suggest that for the items selected for this study, there is a persistent potential for the introduction of bias in the estimation through the application of the current adjustment methodology.

Of the 30 comparisons between the original Wave 1 estimated totals and the corresponding CUR estimates for the selected SIPP items, 18 of the related differences were statistically significant, as denoted by the highlighted table entries. Moreover, for many of these comparisons there were model-based weighting alternatives to the CUR estimates that appeared to have performed better. For example, for the Wave 9 respondents of the 1993 panel the estimated relative difference for poverty is -3.03% for the CUR procedure. This is compared to estimates of 0.46 and 0.67 for the LR and LR/X weighting alternatives. For the Wave 12 respondents of the 1996 panel, the CUR's relative difference for food stamps is -1.57%, which can be compared to the -0.13 associated with the LR/X procedure. Note also that for the Wave 12 respondents of the 1996 panel the relative difference for unemployed for the CUR method is -5.62%, for which the corresponding estimate for the LR/X approach is -0.30%. For the 2001 panel the CUR Wave 9 respondents' estimated relative difference for food stamps was 2.63%. The corresponding estimate for the LR model alternative was 0.34 percent.

There were indications of potentially significant longitudinal nonresponse bias in estimates for the poverty and health insurance items more frequently than for the other items considered. However, in most cases, the alternative estimates for health insurance were not statistically different from the current estimates.

Comparisons were also made, for the race and ethnicity domains, of estimates resulting from the current longitudinal noresponse adjustment with estimates for the selected adjustment alternatives. However, in the interest of conserving space the related tables are not provided in this report. Patterns for the estimates of nonresponse related error were less evident for these groups than those exhibited in the estimates for the item totals. For these domains there appears to be more items with a potential for significant longitudinal nonresponse bias; moreover the magnitude of these biases are expected to be somewhat larger than those of the item totals. The results were rather mixed and varied by domain and panel. Results of the comparisons for the White and Black subgroups varied considerably among the three panels, while results for the Other Races domain seem to suggest a strong potential for longitudinal nonresponse bias for almost all of the selected items. The empirical results for the Hispanic subgroup provided more frequent indications of the preference for the current nonresponse reweighting procedure.

## 5. Discussion

The increased rate of wave nonresponse and attrition in the Survey of Income and Program Participation, and the potential for significant related bias in the estimates for the principal survey items, appear to warrant a continued effort to develop and maintain a longitudinal weight adjustment procedure designed to compensate for nonresponse. The results of the study summarized in this paper provide some insight into the possible extent of nonresponse related error in SIPP estimation, and they permitted some comparisons of plausible response modeling and weighting cell adjustment alternatives. Based on the current evaluation criteria, there is seemingly no single nonresponse weighting cell scheme that is uniformly “better” than the other alternatives considered. However, the results from the application of some of the methodology explored in the study for item totals are encouraging. Weighting cell adjustment procedures, that take advantage of the utility of the historical (previous wave) survey data in the definition of the applicable model, and consequently the nonresponse weight adjustment cells, should warrant strong consideration in the effort to compensate for longitudinal nonresponse.

It is desirable to use longitudinal surveys such as SIPP to study the dynamics of populations and economic phenomena. The study of event histories, spells or durations and transistions provide planners and policymakers with data and relevant results crucial to the development, maintenance and evaluation of major state and national programs. As with the cross sectional estimates, longitudinal nonresponse can have a substantial impact on the quality of derived longitudinal estimates or measures. Another objective of the study of longitudinal nonresponse in SIPP is to investigate its effects and those of related adjustments on longitudinal modeling and estimation. Research is under way, which is designed to identify the theoretical framework and develop valid empirical methods for assessing the relationships between longitudinal nonresponse and duration models for SIPP program participation.

Anticipated future work in this area includes the study of the effectiveness of a class of distance measures as a tool for comparing alternative weighting strategies and the selection of a highly efficient set of nonresponse weighting cells. In addition, efforts will also examine the effect of adjustments for longitudinal nonresponse on duration estimation, survival functions and other event history analyses of SIPP data.

## Acknowledgement

The author expresses his appreciation to Julie Tsay of the U.S. Census Bureau for her adroit and diligent efforts in providing statistical computing support for the study considered in this paper.

## REFERENCES

Da Silva, D. and Opsomer, J. (2004). "Properties of the Weighting Cell Estimator Under A Nonparametric Response Mechanism," Survey Methodology, Vol. 30, No.1, 45-55.

Dufours, J., Gagnon, F., Morin, Y., Renaud, M. and Sarndal, C. (1998 ) "Measuring the Impact of Alternative Weighting Schemes for Longitudinal Data," 1998 Proceedings of the American Statistical Association, Survey Research Methods Section, Alexandria, VA: American Statistical Association: 552-557.

Hendrick, M. R. (1996)." The Creation and Evaluation of Longitudinal Nonrresponse Cells for the 1996 Survey of Income and Program Participation," 1996 Proceedings of the American Statistical Association, Survey Research Methods Section, Alexandria, VA: American Statistical Association : 575-578.

Rizzo, L., Kalton, G., Brick, M. And Petroni, R. (1994). "Adjusting for Panel Nonresponse in the Survey of Income and Program Participation," 1994 Proceedings of the American Statistical Association, Survey Research Methods Section, Alexandria, VA: American Statistical Association: 422 -427.

Sarnal, C.E. Swenson, B. and Wretman, J. (1992). Model Assisted Survey Sampling. New York: Springer-Verlag.

Slud, E. and Bailey, L. (2006). " Estimation of Attrition Biases in SIPP." 2006 Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association: