

DISCUSSION

Constance F. Citro, Nat. Research Council, Michael L. Cohen, Energy Information Admin.
2101 Constitution Ave., NW, Washington, DC 20418

General Remarks

We want to commend the Census Bureau for supporting research on the effects of sample attrition on estimates from the Survey of Income and Program Participation (SIPP) and on methods of developing longitudinal weights that appropriately adjust for attrition. Also, we want to commend the authors for the quality of their papers and their contributions to furthering our understanding in this area.

One cannot overstress the importance of work on adjusting data in SIPP for the effects of attrition from the sample--work on adjusting the data for item nonresponse is also important, but would need a session in and of itself.

The basic problem is that a high proportion of cases in SIPP fail to complete all interview waves. Not only does sample attrition increase the variance of estimates with the remaining cases, but there is the potential for bias to the extent that weighting or other adjustments do not appropriately account for differences between those who complete all waves and those who attrite.

Understanding and compensating for the effects of attrition on SIPP estimates is important because of the many policy and research uses of SIPP, such as estimating participation in government assistance programs and the patterns of such participation (e.g., the length of spells and whether multiple program participation is concurrent or sequential).

A recent Committee on National Statistics panel, chaired by Graham Kalton, recommended that SIPP become the source of the nation's official income and poverty statistics, replacing the March Current Population Survey (CPS). Such a role for SIPP further raises the stakes.

Indeed, comparison of poverty rates from the 1984-91 SIPP panels with March CPS rates consistently shows a pattern whereby the SIPP poverty rate is 2-3

percentage points below the March CPS rate. There are many possible reasons for such a difference, which has important public policy implications were SIPP to be used for poverty estimation. An analysis by Lamas, Tin, and Eargle in a 1994 paper examined several possible explanations and suggested, by a process of elimination, that improvements in the SIPP design and income reporting at the lower end of the income distribution may account for most of the difference; however, they estimated that attrition bias accounted for some fraction, underscoring the need for the kind of research undertaken by the authors at this session.

The planned redesign of SIPP, to be introduced with the 1996 panel, increases the stakes even more, because the proposal is to have four-year panels introduced every 4 years with no overlap across samples. The added length of each panel will facilitate the use of SIPP for important analyses of income and program dynamics, but will further reduce the number of cases available with complete data for all waves.

Cross-sectional estimates from SIPP for, say, a calendar year (e.g., annual poverty rates) can benefit from weighting procedures that use all of the cases with complete information for that year, including cases that subsequently attrite. However, with a 4-year, nonoverlapping design, calendar-year estimates for 2 out of 4 and even 3 out of 4 years will reflect about as high a rate of attrition as full-panel longitudinal estimates (this is because most of the attrition occurs in the first 16 months or so of each panel).

With regard to longitudinal estimation, Hill has demonstrated convincingly and disturbingly that, due to nonrandom attrition, SIPP may underestimate important kinds of transitions and, further, that the currently available weights do not improve matters. Hill looked at marital disruptions but recommended rightly that income, poverty, and program participation transitions be

investigated as well.

The other three papers--by Rizzo et al., Folsom and Witt, and An et al.--make important contributions to understanding the properties of different types of weighting adjustments to attempt to compensate appropriately for sample attrition in SIPP--although none of the authors has identified the magic bullet and, clearly, more work is needed.

In this regard, we encourage the Census Bureau to place the research on longitudinal weighting adjustments in the context of other research on strategies for handling wave nonresponse. The papers presented here focus on strategies for reweighting sample cases with complete information for all panel waves. Other analyses that have looked at patterns of wave nonresponse have identified instances in which cases are missing only one or two waves within a panel. Imputation for the missing waves for these cases could significantly increase the number for analysis and thereby reduce the number for which some type of weighting adjustment is still needed.

Remarks on Specific Papers

We will first remark on each of the individual presentations, and will conclude with some general comments. We will also engage in some statistical name-dropping, in that we are not very knowledgeable about some of the techniques that we will mention, and therefore, there is a greater chance that we are in error in suggesting them.

Hill's paper on "Weighting for Nonresponse in Event History Analysis" presents a terrific model that has the two advantages of agreeing with vital statistics and also having a solid subject-matter interpretation. SURF seems ideal for this problem. To confirm this advance, Hill should: (1) as Hill mentions, try this technology out on other end points, such as change in poverty status, and (2) try this technique on other time periods.

Some (unimportant) questions follow. First, how was the standard deviation of the estimated correlation arrived at? Second, were other models of

correlated competing hazards looked at? Third, does the current method have an E-M algorithm interpretation? If not, could the E-M algorithm be used? And last, was model building of any kind attempted in deciding on which covariates to use, in what form to use them, and what interactions to examine?

Finally, there is a nice confirmation of the approach from the correspondence of the weighting scheme with current practice when the correlation is set equal to 1.

Rizzo, Kalton, Brick, and Petroni's paper, "Adjusting for Panel Nonresponse in the Survey of Income and Program Participation," is extremely well-written, making the job of discussant much easier. It is a very thorough analysis of the problem, providing results from a wide variety of approaches to its solution.

For Rizzo et al., we have the following questions. First, would a binomial model have any hope of succeeding? (We would guess not.) Second, would CART be a useful alternative to CHAID, either for identification of good covariates or estimation of the probability of nonresponse? Third, did the covariates' regression coefficients have the right signs? We think they did with respect to our intuitions, but our intuitions are not that well developed. We add that the covariates used and their signs remind us of census undercoverage models, which suggests looking at those models for help. (One possible covariate is an indicator variable for residence in a major city.) Lastly, why not use both the predicted response rate for individuals and the observed response rate for cells together in an Empirical Bayes approach?

Some final remarks for Rizzo et al. are: (1) given the correlations of the various approaches it seems clear that all the methods are picking up the same structure; (2) it is a great idea to repeat the analysis for another wave; and (3) a priori, we like raking since the interactions--which we believe to be less important--are ignored, and the marginal information is used in a smooth way.

An, Breidt, and Fuller's paper, "Regression Weighting Methods for SIPP Data," makes use of a three-phase structure

that provides great clarity to the entire problem. It allows full use of the available data and relationships at each phase. We have two questions: (1) How can one theoretically choose between the three candidate approaches? and (2) How are the categorical variables, some of which are ordered, some of which are integral, etc., used in the regressions?

We would like to add that we do not believe that the underlying multivariate normal orientation of the approaches is a problem, since what one needs to be well-behaved are ratios of sums, which typically are.

Folsom and Witt's paper, "Testing a New Attrition Nonresponse Adjustment Method for SIPP," presents a very useful alternative to the other approaches. The paper makes use of an excellent idea to use separate models for separate subpopulations. Further, the primary technique is a clever means for reduction of variance inflation.

We have two questions. First, could use of reweighting to CPS control totals reduce the worry about variance inflation? Second, how are the U and L limits chosen? Are they data driven?

We have some concluding remarks. First, to properly compare these approaches, especially the last three, there is a need for a large artificial data simulation study. Second, the question of which approach to use depends to some extent on whether one is addressing a specific use of the data, or whether one is producing a general use data set. Third, there is the interesting question

of global versus local modeling. More effort should be expended on identifying subsets of the population that need to be separately treated. Fourth, there is the interesting question of parametric versus nonparametric modeling. Both approaches seem to have benefits in this problem and should be further investigated. Fifth, the optimal estimator (or better, the chosen estimator since it would be impossible to define optimality), when it is discovered, should have the property of being stable over time. Sixth, we would like to put in a plug for logistic regression diagnostics to help build the logistic regression models of Hill and Folsom/Witt. The primary contributions are Fowlkes and Landwehr-Pregibon-Shoemaker. Finally, why can't one use intermediate wave responses for extra information?

References

Citro, Constance F. and Graham Kalton, eds. (1993) The Future of the Survey of Income and Program Participation. Panel to Evaluate the Survey of Income and Program Participation, Committee on National Statistics, National Research Council. Washington, D.C.: National Academy Press.

Lamas, Enrique, Jan Tin, and Judith Eargle (1994) The Effect of Attrition on Income and Poverty Estimates from the Survey of Income and Program Participation (SIPP). Paper presented at the Conference on Attrition in Longitudinal Surveys. Bureau of the Census, Washington, D.C.