**THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION**

**LONGITUDINAL ATTRITION IN
SURVEY OF INCOME AND PROGRAM
PARTICIPATION (SIPP) AND SURVEY
OF PROGRAM DYNAMICS (SPD)**

No. 242

Denton Vaughan
U. S. Census Bureau
Fritz Scheuren
NORC, University of Chicago

Longitudinal Attrition in SIPP and SPD §

Denton Vaughan, U.S. Census Bureau, and Fritz Scheuren, NORC, University of Chicago
Contact Author denton.r.vaughan@census.gov

Key Words: Nonresponse bias, Social Security earnings records, social security number (SSN) reporting, family and person reweighting.

Abstract: Longitudinal surveys like the Survey of Income Program and Participation (SIPP) and the Survey of Program Dynamics (SPD) suffer from attrition that grows as the survey process is extended out in time. Until recently there has been little research examining directly the effect of attrition on longitudinal change measures. This paper attempts to use Social Security longitudinal earnings records for that purpose.

1. Introduction

The 1996 welfare reform legislation has generated a great deal of new research on the low income population (e.g., [1]). Many useful analyses have relied upon existing surveys like the March Supplement to the Current Population Survey (CPS). Entirely new surveys have been mounted, such as the National Survey of America's Families (NSAF) and the SPD, which have yielded valuable insights as well. Most efforts, CPS and NSAF among them, are essentially cross-sectional in nature. This means that they can, at best, measure only *net* change. This is unduly limiting for some purposes, especially when researchers want to follow particular cohorts of low income people beginning before welfare reform and extending over long periods of time after reform. With a longitudinal survey, *gross* change can be measured, and the SPD was devised specifically to meet this need. Unlike the CPS or NSAF, the SPD is fully longitudinal in its approach.

SPD used as its starting point the 1992 and 1993 incoming panels of the SIPP. Because SIPP respondents had already been surveyed many times before their use in SPD, there was a concern that nonresponse in SPD could be sizeable – and it was. After all, at the end of SIPP, respondents had been led to expect that they would not be interviewed further. While the SPD design was a bold attempt to capture the longitudinal effect of the Welfare Reform Act of 1996, it was criticized because of its high nonresponse rate, which rose, at one point, to 50%.

In general, as nonresponse rate increases, the potential for nonresponse bias also increases. For example, the relationship between the nonresponse rate and nonresponse bias in simple random sampling is of the form:

$$\overline{y}_r = \overline{y}_n + \frac{m}{n}\ (\overline{y}_r - \overline{y}_m),$$

where (m/n) is the nonresponse rate and the averages in "y" reflect values taken from respondents (subscript "r"), the ideal average values on the entire sample (subscript "n") the average from the unobservable nonrespondents (subscript "m"). The expression above makes clear why researchers were concerned that there might be serious bias in SPD measures of the effects of welfare reform (since m/n was so large). Another look at the expression reveals a second key to bias: how different respondents are from nonrespondents. If they differ little, then bias will be minimal, even with a poor response. This is the area of research this paper will focus on – examining longitudinal differences between respondents and nonrespondents.

An important advantage that SPD has over other welfare reform data sets is that it can be matched to the earnings records maintained by the Social Security Administration (SSA). A longitudinal survey matched to a longitudinal set of administrative records offers an ideal framework for studying the impact of welfare reform. Until now, though, the SSA earnings records have not been used in this way to any significant degree. This paper is among the first to change that (for another see for example, [2]).

Organizationally, the paper is divided into eight brief sections, beginning with this introduction (Section 1), followed by a section on the specific goals and objectives of the study (Section 2). The matching of the survey and SSA earnings data is described next (Section 3). Reweighting methods are then detailed in Section 4. In Section 5, initial comparisons are made between those matched cases who attrit out of the survey and those who continue. Section 6 provides additional longitudinal analyses. The paper ends with a few ideas on future research (Section 7), followed by a summary of the conclusions we have so far (Section 8).

2. Objectives of Study

The present SPD nonresponse study has three main

objectives: (1) to ready the matched SPD and SSA data sets for longitudinal analysis, (2) to look at the potential longitudinal bias introduced by attrition, and finally, (3) to use the administrative data to adjust any longitudinal biases observed, to the extent possible. The key difference in the present study versus most of the earlier match studies conducted by the SSA and the Census Bureau over the last several decades is the emphasis on *longitudinal* impacts. The change over time in a person's circumstances or those of his or her family is the topic of most interest here, especially the effects of the welfare reform legislation felt after 1996.

Longitudinal surveys (like SPD) inherently suffer from growing nonresponse. Such studies can still be competitive over independent or nearly independent repeated cross-section designs (like NSAF) for measuring *net* change, if adequate adjustments are made for the greater nonresponse. Most of the survey literature, including previous analyses of SPD, has emphasized this point.

Gross change measures cannot be estimated at all in a survey like NSAF. Such measures are only available in undertakings like SPD. Arguably, the measurement of *gross* change is the primary reason for a longitudinal data collection. Surprisingly, an effort to see how well SPD measures *net* change over time, after adjusting for the presence of attrition, has not yet been attempted. The present study remedies this gap focusing on both *net* and *gross* change measurement in the presence of nonresponse. The presence of the SSA longitudinal earnings data makes this possible, since that data are available for both those who attrit after the first interview and those who do not.

3. Matching SPD and SSA

For SIPP adults 15 years or older – the group of interest in this study – social security numbers (SSNs) were reported or otherwise obtained for about 88% of the (unweighted) sample cases in both SIPP years. SSNs were not the only identifiers available, but they were considered indispensable in linking SSA earnings to SIPP interviews. To be considered usable for matching, however, not only did the case require a valid SSN but there had to be a high level of agreement between the survey and administrative data on key confirmatory items, mainly names and dates of birth. If SSA and SIPP records matched on SSN but were not close on these other items, then the case was not treated as having been validly linked. We also elected to exclude a small number of other SIPP interviews because of data problems. After these exclusions we were able to examine, on a weighted basis, about 80% of the wave 1 SIPP interviews in both 1992 and 1993.

A matching bias could be present because not all SSNs were available. In particular, by shifting to the SIPP sample with usable SSNs we could be simply substituting concerns about one form of bias for concerns about

another. While in theory a possible matching bias may be large, the evidence suggests otherwise. In other Census Bureau survey settings (notably the CPS) where SSNs have long been used to match administrative data, once a modest amount of adjustment is done there is virtually no residual bias evident [6]. To check whether this was also true for SIPP, an extensive set of sample person tables were produced that looked at differential reporting rates. In all, 20 SIPP comparisons were produced for different variables or combinations of variables. In most cases there were only tiny differentials, seldom more that a percentage point.

4. Person Reweighting of SPD

The review of the SSN missingness patterns in SIPP and our earlier experiences suggested there would still be advantages to reweighting the matched cases to bring them back up to the original SIPP weighted controls. This was done using the weighted results for the linked cases from SIPP, versus all cases. The Deming-Stephan raking algorithm [3] was employed to align a subset of the dimensions examined, so that the original and matched SIPP files agreed. The dimensions (and associated levels) we forced agreement on were age (2 levels only), household relationship (3 levels), race and ethnicity (3 levels), family monthly income (2 levels), family welfare ratio (2 levels), and citizenship (2 levels). The raking was done separately for the 1992 and 1993 SIPP data.

A much more extensive adjustment was possible, but we feel that for national analyses the results can be used as initially reweighted. In making this assessment we made a great many comparisons between the full or original SIPP samples and the reweighted cases. The variables we looked at were gender, ethnicity, age, household status, welfare ratio and citizenship, since these were partially controlled, plus family size, the presence of welfare income, family earnings, family annual income, and educational status, which, while not directly controlled, were viewed as very important. For all of these variables, the agreement at the national level between the original and matched results was good to very good.

An important feature of raking is that after adjusting the weights and retabulating of the data, as above, we can continue the adjustment from where it was left off, possibly using entirely new dimensions (e.g., region) or the same dimensions but new combinations of levels. After we complete the rest of our preliminary analyses, this is almost certainly what we will want to do. For now, however, we are ready (in the next section) to go on to the most important step, of comparing linked SSA earnings data for attriters and nonattriters separately.

5. Initial SSA Earnings Comparisons

Typically, in handling nonresponse in longitudinal surveys, the attrition-truncated sample is reweighted to account for the differential characteristics of attriters –

ordinarily, as observed, before they stop participating in the survey. Arguably, this approach assumes that the characteristics of attriters, as observed while participating in the survey, remain fixed after they leave or, like age, change in a predictable way. But what about behavioral variables, like employment or earnings? Are the shortfalls seen in these characteristics permanent or can they change? Is there, for example, a regression to the mean effect?

In the current study we can assess questions like these using the longitudinal SSA Summary Earnings Records of both those people who continue in the survey and those who leave, since we are able to track earnings of both groups independently of their interview status in later waves. This is possible, as discussed above, because the survey obtained and validated Social Security Numbers (SSNs) for sample members at the first SIPP interview. Consequently, missingness due to sample loss after the first interview is avoided for SSA earnings.

While both the 1992 and 1993 SIPP were used as starting points for the SPD, so far we have only examined the earnings data for the 1993 SIPP. For confidentiality reasons, the 1993 results we have can only be shown in summary fashion. This has been done in Table 1, where the original 1993 SIPP sample members were divided in two groups based on their attrition status as of the last interview of the SIPP: those who continued in the survey (i.e., persons interviewed in both waves 1 and 9 of the 1993 SIPP panel, or "Continuers") and those who left the survey (i.e., were interviewed in wave 1 but not wave 9 of the 1993 SIPP, or "Attriters").

Table 1 examines annual earnings at two points in time over the 7-year period: 1992 (the year of the initial SIPP interview) and 1999 (the fourth year covered by the SPD, and the last year covered by our extract from the SSA earnings records). We found in both years that the median earnings of all attriters was roughly half that of those continuing (based on all people, including zero earners). In a general way, this confirms prior research based on the observable survey characteristics (e.g., as shown in several of the papers in [4]). While nominally, the attriters appear to become more selective over the period, a difference of difference of test does not sustain this finding. For those with positive earnings, though, a somewhat different pattern may be present. Selectivity is less evident in the base year, and the gap between attriters and those who continue does narrow (as indicated by a difference of difference test which shows a statistically significant convergence between median earnings of the two groups over the period.) However, all differences in medians remained statistically significant.

In some ways, this comparison is unfair. If we delete those who left the survey universe (i.e., by eliminating those who attrited because of death, institutionalization, moving overseas or into Armed Forces barracks), then a pattern of convergence become evident. Not only are

attriters less selective to begin; they become more like those who continued over time. In fact, by 1999 the difference in median earnings is not statistically different. The decline in selectivity also appears to be less marked among those with positive earnings. While the gap in median earnings also nominally declines, the decline is not statistically significant. What do these patterns reveal?

Table 1: Median Annual SSA Earnings, the Year of Initial Interview (1992) and Last Year Observed (1999)

(Persons 18 years of age or older; medians in 1993 dollars)

Participation Status	1992	1999
Total persons. age 18 or older Continuers	$7,500	$8,700
Attriters	3,800*	3,900*
Attritors as % of continuers	51	45
With Positive Earnings† Continuers	18,000	22,000
Attriters	13,000*	19,000*
Attritors as % of continuers	74	84
Persons, age 18 or older *† (Excluding Universe Leavers) Continuers	7,500	8,700
Attriters	6,400*	8,900
Attritors as % of continuers	85	102
With Positive Earnings (Excluding Universe Leavers) Continuers	18,000	22,000
Attriters	14,000*	19,000*
Attritors as % of continuers	76	85

*Difference by year is significant at the 90% confidence level.
† Difference of differences between years is significant at the 90% confidence level.

Source: 1993 SIPP matched to SSA earnings records.

There are three main points, which stand out with regard to *net* and *gross* change measures, with some anticipated, some not. First, attriters from the survey are likely to have lower earnings at the time of the first SIPP interview than those that continue. While this point is not new, it is now confirmed by the matching of SSA earnings to the attriters and nonattriters. The next two points are more interesting and arguably controversial. Surprisingly, the initial observed gap in earnings is not permanent. Relying on the SSA earnings records, we find a

"regression-to-the-mean effect" that brings the two groups back together in about 7 years. In fact, once decedents are eliminated, overall differences in earnings are not statistically significant bewteen attriters and continuers.

It is not obvious on its face that the gap would *not* be permanent. Few studies [5] have even pointed this out before and not with this kind of evidence. Speaking technically we are able, with the matched data, to look at the extent that the missingness is ignorable or nonignorable *longitudinally* for variables related to SSA earnings measures. And we found the missingness is *largely ignorable* for *net* change and *largely nonignorable* for *gross* change measures.

Why are the results largely ignorable for the *net* change measures examined? One of the main reasons was our ability to eliminate decedents; this small step has a major effect on the estimates. So important is it that the Census Bureau should consider routinely matching SIPP and other continuing longitudinal household surveys to the National Death Index, even when it does not match such surveys to other administrative records. The "death clearance process," we relied on in using the survey is almost certainly less satisfactory because deaths may well be understated.

Why does it take just a few years for the *net* change measures to become similar? Is this just a detail or does it have some importance? The analyses done in the next section seem to shed some light on the issue.

6. Further Earnings Comparisons

We do know from unrelated work done some time ago [6] that major personal "shocks" (death, divorce, job loss, even moving) can take several years to work themselves out. In the interval, the event seems to act like a gravity well and distorts measurement. Could this be what is happening when we have nonresponse? Could nonresponse be a "marker" for these other events?

This lagged response, if borne out, could be very important for those using the SPD. Our research so far has focused on the attrition that took place by the end of SIPP. We have yet to look at the attrition that occurred in SPD itself, where the nonresponse event is much closer in time to the year 1999. Conceivably, we will find biases in both *net* and *gross* change measures from this later nonresponse, if unadjusted.

The implications of these results are important, at least in suggesting working hypotheses that need further testing. How might we test our understanding further? If nonresponse was a marker for a life shock, we should find that SSA earnings before the survey would be more similar for attriters and continuers than they were at and in the initial years after the survey. This requires another look at the SSA earnings data and the pattern of earnings *before* as well as *after* the survey. Table 2, shown at the end of this paper, summarizes what we found.

The life shock hypothesis is not borne out. Instead, we find fairly smooth patterns that do not have any marked discontinuities around the interview year. The difference in the slope of change in median earnings over time for attriters and continuers seems to be the dominant difference between the two groups.

7. Possible Next Steps

More work is required to understand the missingness patterns in SPD to which the SSA administrative data give us insight. This is certainly true if we are to achieve our third goal in the study, which is to use the administrative data to adjust for any longitudinal biases observed. Below is a discussion of some of the remaining methodological work, and of the possible benefits to welfare reform analysis of an enhanced SPD.

Further research is warranted to confirm and deepen the present findings. A more thorough adjustment for missing SSNs may even be in order. For example, most of the missing SSNs arise because of refusals and if confidentiality hurdles could be overcome the possible effects of refusals could be directly investigated. Indeed, this possibility was part of the original 1973 Exact Match Study [7]. It was not pursued in the end because at that time most of the SSN missingness in the CPS was due to other causes. Now this is no longer true – a distinct shift that parallels other Census Bureau experiences in nonresponse over the nearly 30 years since then. One of the reviewers of an earlier draft of this work noted that attriters had a higher SSN missing rate than nonattriters. This does not directly affect the results given here, but needs to be taken account of in providing an overall set of reweighted results.[1]

Examples of other additional research include expanding the analysis to include the 1992 SIPP panel (which was the starting point for the other half of the SPD sample), incorporating results of the attrition experienced by the actual SPD sample, including the efforts at re-contact and incentives to continue or resume participation in the survey. Finally, it would be valuable to extend the years of SSA earnings data to fully cover the entire SPD reference period (through 2001).

While there is a lot of work involved, we feel fairly comfortable that improving SPD for use in measuring *net* change may be straightforward, given what we now know, at least for persons with earnings. We also expect to be able to extend that work to the more difficult problem of adjusting families with SSA earners [8]. Both of these goals were ones we anticipated accomplishing when we started out.

As regards improving inferences concerning *gross* change measures in SPD, our research is only at the

---

[1]The missing rate *was*, in fact, higher for those who eventually attrited (16 percent) compared to those who continued in the survey (10 percent).

beginning. While methods in the statistical literature exist that could be applied to *simultaneously* adjusting for both *net* and *gross* change, the results are unclear. Clearly, though, the value of the matched longitudinal survey and administrative data as a tool for longitudinal survey analyses is now emerging as even more important.

Both SPD and NSAF have great value in welfare reform analyses, especially if used together – something not done to any extent yet. NSAF, for example, has excellent data for 13 selected states, something SPD cannot duplicate. Still, marginal analysis resources might best be spent with SPD if the interest is primarily in *net* national change measures, because SPD has so many administrative covariates to adjust for bias. And, of course, to measure *gross* national changes arising from welfare reform, SPD stands alone. But it must be heavily adjusted using available administrative data before it can be used safely for this purpose.

8. Conclusions Thus Far

From the analyses done thus far, we conclude that eventually *net* change measures from SPD, with only modest adjustment, may be completely comparable to cross-section surveys, like CPS, for earnings and variables closely related to earnings, such as work experience.

On the other hand, for SPD to be used successfully to measure *gross* change, the data adjustment is likely to be greater. The very success of SPD as a measure of *net* change creates this seeming paradox. The paradox arises because those who attrit start out with lower incomes than those who remain in the survey longer. For attriters to catch up in about seven years, therefore, they necessarily have to have a different time trajectory than the more stable group who remain. Only the presence of the administrative data makes this phenomenon noticeable and provides the opportunity to adjust for it.

At the root of this discussion is one of the major poverty research issues: the degree to which poor and low income individuals escape permanently from their circumstances and, in the context of welfare reform, the extent to which policy interventions increase (or decrease) those chances. Clearly some of the poor and near poor do escape but many (especially women) churn most of their lives, rising and falling back into the low income population, over and over again. The Michigan Panel Study of Income Dynamics (PSID) provided this insight some time ago.

Low income individuals do not simply either increase their income or stay at the same income level. Many can and most do go both up and down. Income variability for low income individuals is the rule, rather than the exception (as it is for the middle class). Whether this churning phenomena will continue throughout the era of welfare reform is still unknown. For SPD to be used to find out, however, it seems clear that some further

adjustment of the survey will be needed, otherwise *gross* change measures will be biased downward.

Before ending this paper, the phrase "regression to the mean" as used above deserves a separate comment. In our view, the main research question is to what extent are attriters part of the same distribution as continuers and hence subject the regression to the mean effect. Or to what extent, as in the case of decedents, they are not in the same distribution at all.

To be able to say that attriters are part of the same distribution is equivalent, it might be noted, to making the statement that there being nonrespondents was ignorable after adjustment. And for *net (*but not *gross)* change measures this regression to the mean effect does seem to be present after six years.

That we observe some regression to the mean, after a period of years and after adjusting for decedents, suggests, in part, that we have found evidence that the missingness of *net* change measures is nearly ignorable if we go out far enough in time. For *gross* change measures, however, much work remains.

References
[1] National Research Council (2001). *Evaluating Welfare Reform in an Era of Transition*. National Academy Press: Washington.
[2] Huynh, M., Rupp, K., and Sears, J. (2001). **The assessment of Survey of Income and Program Participation (SIPP) benefit data using longitudinal administrative records.** *http://www.FCSM.gov.*
[3] Deming, W.E. and Stephan, F. F. (1940). "On a Least Squares Adjustment of a Sample Frequency Table When the Expected Marginal Totals are Known," *Annals of Mathematical Statistics.*
[4] Special Issue on Longitudinal Surveys, *Journal of Human Resources*.
[5] Fitzgerald, J., Gottshalk, and Moffitt, R. (2001). An Analysis of Sample Attrition in Panel Data. Journal of Human *Resources*.
[6] Scheuren, F. and Oh, H.L. (1980). Report No 11 in the *Interagency Data Linkages Series*. Social Security Administration.
[7] Herriot, R. and Scheuren, F. (1980). "The 1973 Exact Match Study," reprinted in Report No 11 in the *Interagency Data Linkages Series*. Social Security Administration.

[8] There are many ways to adjust family weights to fit totals, as detailed in a special 1987 issue of *Survey* *Methodology* that covers this and related topics. The paper by Chip Alexander is especially important.

Table 2.---Annual median earnings from the Summary Earnings Record (SER) by Universe Membership and Survey Participation as of Last SIPP Interview, Selected Years [constant dollars of 1993]				
Number of Years or after the initial SIPP interview	Total	Continued [1]	Attrited [2]	Earnings of attritors as % of continued
All Sample Members Interviewed in Wave One, age 18 or older				
Before [4]				
Seven ('85)	$3,200	$4,300	$ 600	14
Four ('88)	5,600	6,700	2,500	37
Three ('89)	6,100	7,300	3,200	44
Two ('90)	6,400	7,400	3,700	50
One ('91)	6,300	7,300	3,400	46
Year of 1st First Inter- view ('92)	6,500	7,500	3,800	51
After [4]				
One ('93)	6,900	7,800	4,300	55
Two ('94)	7,300	8,300	4,500	54
Three ('95)	7,500	8,600	4,300	50
Four ('96)	7,600	8,600	3,900	45
Seven ('99)	7,700	8,700	3,900	45
Excluding Universe Leavers Who Exit Sample Subsequent [3] to Wave One, age 18 or older				
Before [4]				
Seven ('85)	3,700	4,300	1,90 0	44
Four ('88)	6,200	6,700	4,300	64
Three ('89)	6,800	7,300	4,300	59
Two ('90)	7,100	7,400	5,900	80
One ('91)	7,000	7,300	5,700	78
Year of 1st First Inter- view ('92)	7,300	7,500	6,400	85
After [4]				
One ('93)	7,600	7,800	7,000	90
Two ('94)	8,100	8,300	7,600	92
Three ('95)	8,400	8,600	8,000	93
Four ('96)	8,500	8,600	8,200	95
Seven ('99)	8,800	8,700	8,900	102

Notes to table 2

Note: Medians rounded to two significant digits to protect against disclosure. Tests of statistical of significance are pending.

[1] Interviewed in the first and last interview of the SIPP.

[2] Interviewed in the first interview of SIPP; noninterview in the last SIPP interview.

[3] Universe leavers include those who exit the survey because of death, institutionalization, or moving overseas or to an armed forced barracks, as of last contact. Status measured as of last completed interview.

[4] Calendar year of interview and annual earnings in brackets.

Source: 1993 Panel of the Survey of Income and Program Participation (SIPP) exact matched to Social Security Administration Summary Earnings Records. Cases with usable social security numbers reweighted to represent survey population as of Wave 1 of the survey.