

**THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION**

**INCOME POVERTY TIMES SERIES
DATA FROM THE SURVEY OF INCOME
AND PROGRAM PARTICIPATION**

No. 207

**V. J. Huggins and F. Winters
Bureau of the Census**

**THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION**

**INCOME POVERTY TIMES SERIES
DATA FROM THE SURVEY OF INCOME
AND PROGRAM PARTICIPATION**

No. 207

**V. J. Huggins and F. Winters
Bureau of the Census**

INCOME AND POVERTY TIME SERIES DATA FROM THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Vicki J. Huggins and Franklin Winters, U.S. Bureau of the Census¹
Vicki J. Huggins, U.S. Bureau of the Census, Washington, DC 20233

I. INTRODUCTION

The Survey of Income and Program Participation (SIPP) is a complex longitudinal survey conducted by the Census Bureau to provide information for U.S. policy makers and academic researchers on topics such as poverty, government program participation and eligibility, health insurance coverage and income distributions. Beginning in 1996, the SIPP will go into the field with a redesigned sample based on 1990 Decennial Census information, with oversampling for the low income population and with computer assisted personal interviewing (CAPI). Longitudinal samples will cover 4 years and have an abutting design rather than the current 2 2/3 year overlap design.

The current SIPP overlap design reduces the effects of panel biases for cross-sectional estimation. However, the current longitudinal samples are not large enough to measure important phenomena such as long-term spells of poverty. During redesign planning the Census Bureau determined that longitudinal goals are the most critical for the SIPP program. As a result, a non-overlapping design of larger, longer panels was approved for SIPP redesigned samples.

Among the most important measures from the survey are income and poverty statistics which can be tabulated as a annual time series. Recently, the Census Bureau observed some important phenomena in current poverty and low income data series from the SIPP that is cause for concern. Specifically, there is a consistent drop in poverty and some income estimates across panels from the first to second interview that is larger than expected. Also, there is a consistent pattern of decrease in poverty over the life of a panel - approximately

8 interviews covering 32 reference months currently. These phenomena in and of themselves are troubling, but it becomes even more troubling when the time series carries over from the end of one 4-year panel sample to the beginning of another 4-year panel sample. With the observed decline in poverty estimates over the life of a panel and the higher level reporting at the first interview of a new panel, the jump in the time series resulting from switching to a new panel could be substantial.

This paper will present (1) what we can currently provide for time series data from the SIPP, (2) research in progress to identify the levels of measurement error, attrition, and time-in-sample bias that initially appear to be the root causes of the poverty and income time series phenomena and (3) conclusions about producing official income and poverty times series estimates from the redesigned SIPP.

II. WHAT DOES CURRENT SIPP TIME SERIES DATA LOOK LIKE?

A recent SIPP report by Paul Ryscavage at the U.S. Census Bureau on the economic Health of American Households provides valuable information about the reliability of time series. Figure A provides a startling picture. Quarterly estimates of the number of households with low monthly income (below the poverty threshold) follow virtually the same pattern for all SIPP panels. Estimates drop significantly from Quarter 1 to Quarter 2 (wave 1 interview to wave 1/2 interview) in all panels, even the 1990 and 1991 panels which were fielded during the recession. Separate panel estimates for the same time period differed significantly one-third of the time. This implies a difference effect outside of

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 U.S. Census Bureau.

normal sampling variance and suggest a panel bias problem.

In research conducted by McCormick, Lepkowski and Chakrabarty on the SIPP (1994, 1992, 1988 respectively), very few estimates were found to be affected by time in sample bias and/or attrition. Unfortunately, poverty estimates were among the few examined in McCormick that appeared to be adversely affected. [McCormick did not attempt to separate attrition bias from other time in sample biases.]

Figure B presents combined panel quarterly estimates of poverty households for the same time periods and shows that the quarter 1 to quarter 2 (Q1 to Q2) effect is dampened but not corrected with the current SIPP overlap design.

Panel bias doesn't appear to be as serious for estimates of real median income. There is a pattern of increase in Q1 to Q2 estimates in all but the 1989 panel although they are not significantly different. Figure C provides the combined panel estimates for median monthly income, which also reflects the Q1 to Q2 phenomena, but looks more reasonable than poverty as a time series.

The above discussion focused on cross-sectional quarterly estimates. There is also evidence of similar phenomena in longitudinal estimates. Calendar year poverty rates are provided graphically in figure D. The estimates of the number of poor and poverty rates are statistically different between the 1985 and 1986 panels for the 1986 calendar year. The second calendar year estimate for number of poor is always lower than the first calendar year estimate from the same panel, the 1990 panel as the only exception. Only the calendar year poverty rates from the 1986 panel (CY 1986 and CY 1987) are statistically different from each other though.

We found similar results when we examined poverty estimates by subgroup characteristics. Table 1 provides first quarter poverty estimates for 1991 from the 1990 and 1991 cross-sectional files.

Smaller estimates are consistently observed from the 1990 panel data, i.e., the panel with greater nonresponse by Q1 1991. The most notable result is that the estimates for Blacks do

not follow the trend of all the other estimates. We would like to see if this same result is found when comparing other panels for the same time periods.

One major concern with the SIPP redesign consisting of nonoverlapping consecutive 4-year panels is the possible break in time series from one panel to the next. We simulated what such breaks might look like for both median income and the number of low income households. We forecasted the time series from the 1984 panel for two years to the beginning of the 1988 panel. The break for median income from Q4, 1987 to Q1 1988 is not statistically different. We also plotted the time series using actual combined panel data. Both approaches provide a reasonable time series picture for median income.

Time series simulations for low income households seriously concern us. Simply extending the 1984 panel data from 2 to 4 years using a simple model result in a very large time series break at the beginning of the 1988 panel. Even using true combined panel data to extend the time series leads to a statistically and analytically serious time series break - figure E. Similar results were observed when we repeated the simulations using the 1985 and 1986 panels.

More complicated forecasting models can be devised to predict the break, but the bottom line is that even the real SIPP data from overlap samples result in a serious break. There is no reason to expect this to improve with 4-year nonoverlapping panels.

III. RESULTS OF RESEARCH TO IDENTIFY THE ERROR COMPONENTS

We set out in this research to investigate and eliminate the causes for the Wave 1/Wave 2 phenomena which produces substantial differences in estimates between waves 1 and 2 that are not observed in waves of other panels covering the same reference period and, determine what effects nonresponse has on income and poverty estimates over the life of a panel. We hypothesized that the root causes might include nonresponse/attrition bias, seasonal effects, and reporting error such as telescoping or time-in-sample bias.

more generally, reduce nonresponse bias in SIPP estimates.

1. Nonresponse Surveys

Investigate the conduct of a nonresponse survey to collect information on nonrespondents to use in weighting. The survey could be conducted routinely with routine noninterview adjustments based on the survey data, or the survey could be a one-time or intermittent operation to develop and update parameters for use in adjusting micro or macro data routinely. The survey should focus on collecting information on income to improve income and poverty estimates.

2. Drop Wave 1 Data

Consider dropping the collected wave 1 data and carrying a previous panel over one more wave to bridge the gap and continue the data series. For example, in 2000, we would collect wave 1 data but use a 13th wave from the 1996 panel to compute time series statistics. The next report would come from wave 2 of the 2000 panel. This would probably be a reasonable approach for say median income statistics, but a 13th wave of interviewing is probably not a good bridge for poverty statistics.

3. Alternative Designs/Supplementary Panels

Investigate use of a smaller bridge (overlapping) panel that is fielded around the end of one panel and the beginning of another panel to help smooth the data series. Data from the panel could be used to compute combined panel estimates. This data could also be used for further investigation into attrition and time in sample effects.

We could field small, overlap panels, similar to the current design. The data could be used for modeling purposes to help reduce attrition effects versus simply combining the data with the primary redesigned panel.

These alternative design options could help with the reliability of median income time series statistics, but simply adding an overlap component into the poverty statistics will probably not correct the problems we've seen, unless the overlap panels are introduced more frequently than once a year.

4. Recontact Nonrespondents

During the last interview that overlaps with the start of a new panel, return to all panel nonrespondents and collect information for the bridge time period. In 1996, this will be the end of the 1993 panel.

5. Improving Nonresponse Followup

Continue to investigate procedures to improve nonresponse followup such as use of incentives or shorter questionnaires for hard-core nonrespondents. Tailoring questionnaires, and adding questions that might help adjust for nonresponse. For example, if a respondent provides information that he might move in the next 4 months because of a better paying job, that would be useful information in weighting if he were to become a nonrespondent.

6. Imputation

Continue improving imputation methods for item and person nonresponse, making more use of carry-over methods to improve cross-sectional estimates and imputing for more than 1 wave of missing data when they're bounded by completed interviews. Also, investigate imputation for missing interviews that are at the end of a longitudinal period, say the end of a calendar year or the panel. All of these procedures make use of the fact that we have a substantial amount of data collected for people and even if they don't respond in all waves, making use of their data should be better than weighting others up for them.

V. CONCLUSIONS

Can we provide reliable income and poverty time series data from the redesigned SIPP? Not without some major improvements. The Census Bureau is taking a hard look at this issue. A nonresponse team was set up to develop a profile of nonresponse to help direct future nonresponse research, and many SIPP staff are considering some of the alternatives presented in the previous section.

We recommend that some of the questions about SIPP measurement error be researched through cognitive lab techniques, reinterviews and record checks. Measurement error exists in the SIPP and may be significant for many estimates but nonresponse bias appears to be a much more serious error for poverty estimates.

Thus far, results point to nonresponse as the primary cause for the troubling phenomena we observe.

HIGHLIGHTS

• There does not appear to be any significant external telescoping in wave 1. For example, we compared W4 of the 1990 panel to wave 1 of the 1991 panel to see if there was a higher level of reporting by reference month in W1. The pattern of monthly median income estimates within wave 1 were very similar to the patterns for wave 4 for the same time period but different panel. With external telescoping, we would expect to see much higher reporting at wave 1 in the first month and a larger difference between months than we observed in this data. Also, a record check study conducted in 1984 did not suggest external telescoping to be a problem in SIPP data. More recent record check data should be explored to see if this finding is consistent.

• There is no substantial difference in reporting between waves 1 and 2 that would cause the decline in poverty we observe between Q1 and Q2 estimates.

We compared a distribution of Q1 to Q2 estimates of low income HHs for the same time period but different panels. The distributional patterns are very similar.

• There definitely appears to be a seasonality effect. Quarter one estimates were generally higher than that of the other quarters of the year, regardless of the wave/panel observed. We are in the process of sorting out the seasonality effects and plan to reproduce the time series simulations presented earlier without the seasonal effects.

• There is substantial evidence that nonresponse has an effect on the estimate of number of low income households. The majority of Q1 estimates for the same time period but different panels are significantly different. Also, past weighting research found that we do not fully compensate for the number of low monthly income households in cross-sectional weighting. This underestimation accumulates over time as nonresponse accumulates.

Nonresponse is approximately 7% at the first interview and increases to approximately 22% by the 8th interview. These rates have

increased in later panels. Nonresponse as of wave 8 in the 1992 panel was 24%. At one time, our best guess was that by wave 12 in redesigned panels, nonresponse would be around 25%. If we do not make improvements to current field procedures in redesigned samples, nonresponse could be between 30% and 35%, because we do not know how people will react to 4 years of survey burden compared to 2 2/3.]

To further investigate possible nonresponse effects, we examined income-to-poverty ratios to assess transition patterns in and out of poverty between quarter 1 and quarter 2 estimates for attritors and those who stay in the sample.

Attritors are people who leave by wave 4 of the 1990 panel. They are found in the 2nd column of table 2. Stayers are those persons who responded in waves 1-4. We found that the distribution of transitions is statistically different for all groups, except for households at or below 75% of the poverty threshold in Q1. We conclude that

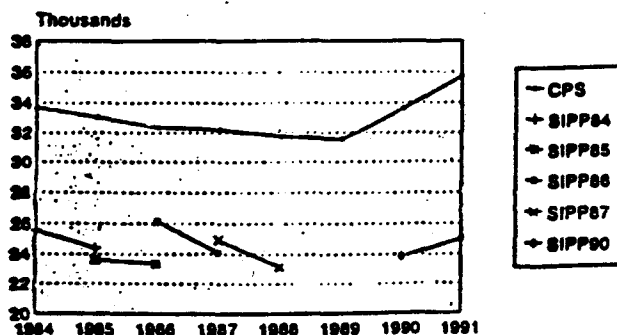
- SIPP keeps those in Q2 that are chronic poverty in Q1.
- By Q2, we lose below borderline poverty when they move completely out of poverty.
- By Q2, we lose above borderline poverty when they move into poverty.
- By Q2, we lose those above poverty when they move into poverty.

The 1991 Q1 poverty rate of only the attritors is 15.7% compared to the complete sample poverty estimate of 11.5%. Distributions of nonrespondents after leaving the survey could vary even more than what is observed here which could result in an even larger nonresponse bias.

We cannot ignore these findings - especially since SIPP is being considered as an official reporter of national level income and poverty statistics. Both cross-sectional and longitudinal estimates are affected. Nonresponse bias appears to be the primary cause.

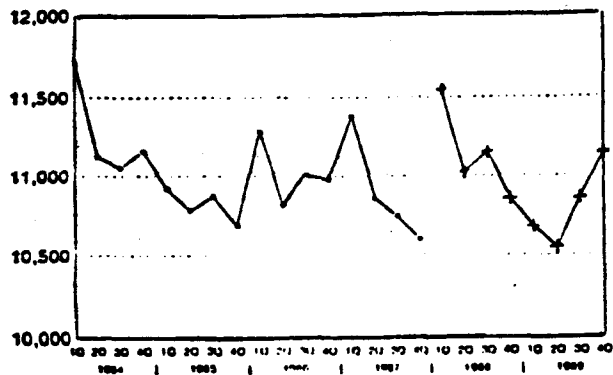
Current measures to reduce nonresponse in the field or minor weighting adjustments for nonresponse bias we have introduced for redesign, are unlikely to adequately correct for the problems we have observed for poverty time series data. Below are some approaches we can consider to improve time series estimates and

FIGURE D: NUMBER OF POOR
Calendar Year Estimates



number in thousands

FIGURE E: SIMULATION OF 4-YEAR SIPP PANEL BREAKS
NUMBER OF HOUSEHOLDS WITH LOW MONTHLY INCOMES - 1986/1987 COMBINED PANEL DATA



- Series 1 + Series 2
IN THOUSANDS

Table 2. 1990 PANEL Q1 TO Q2 POVERTY TRANSITIONS

Table 1. Q1 1991: POVERTY ESTIMATES: 1990 AND 1991 PANELS

CHARACTERISTICS	PERCENT OF PERSONS IN POVERTY	
	1990	1991
All persons	11.4	11.9
AGE		
Under 25 years old	9.0	9.8
25 to 44 years old	4.1	4.5
45 to 64 years old	3.5	3.6
65 years old and over	3.9	4.1
RACE AND SPANISH ORIGIN		
White	3.4	3.6
Black	9.6	9.5
Hispanic Origin	8.2	8.6
Not of Hispanic Origin	3.8	4.1
TENURE		
Owned or being bought by someone	2.4	2.5
Rented for cash	7.2	7.9
Occupied without payment of cash rent	8.2	11.1

Percent of Poverty Level	Percent of People		
	Q1 - Q2	Attriters	Stayers
<75% - 75%		73.7	75.2
- 75% to 95%		11.0	11.8
- 100% to 124%		5.2	5.0
- ≥125%		10.1	8.0
75% to 99%*			
- 75%		19.4	18.6
- 75% to 95%		40.7	48.2
- 100% to 124%		13.6	16.5
- ≥125%		26.4	16.7
100% to 124%*			
- <75%		10.7	6.9
- 75% to 95%		17.7	9.3
- 100% to 124%		29.5	46.6
- ≥125%		43.1	37.3
≥124%*			
- <75%		1.6	.9
- 75% to 95%		1.1	.8
- 100% to 124%		2.6	1.6
- ≥125%		94.7	96.7

* Distributions are statistically different.

We recommend that sample loss be monitored **CLOSELY** by type of nonresponse and characteristics during the initial phase-in of the 1996 panel. It is possible through a redesigned SIPP questionnaire that the profile of nonresponse we see in redesign may be different than the current profile of nonresponse.

REFERENCES

Chakrabarty, R., (1988), "The Effects of Time-in-Sample Bias on SIPP Estimates. A Preliminary Report", Internal Census Bureau Memorandum.

Lepkowski, J., S. Pennell, (1992), "Panel Conditioning Effects in the SIPP", Proceedings from the Survey Research Methods Section, American Statistical Association, pp. 556 - 571.

Marquis, K. and J. Moore (1988), "Using Administrative Record Data to Describe SIPP Response Errors", Proceedings from the Survey Research Methods Section, American Statistical Association, pp. 360 -365.

McCormick, Maryann K., (1994), "SIPP 1984-1988 Panel: Time-in-Sample Effect Study Results", Internal Census Bureau memorandum for the Record.

Ryscavage, P., (1994), Monitoring the Economic Health of American Households With Data From the SIPP, Census Bureau P-70 Report, No. 35.

Short, K., J. Coder and R. Benefield, (1993), "Comparisons of SIPP and March CPS Poverty Estimates", Internal Census Bureau Memorandum to Daniel H. Weinberg.

Fig. A. Households With Low Monthly Incomes From SIPP Panels
(First Quarter, 1984 to Third Quarter, 1991)

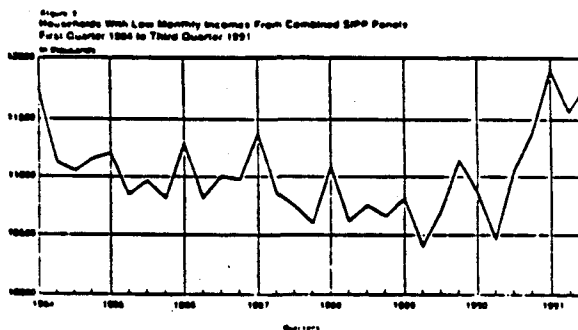
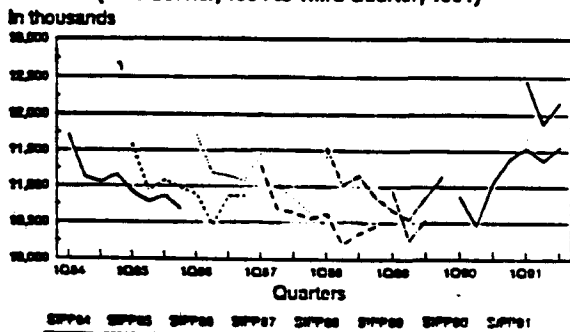


Figure C.
Real Median Monthly Household Income From SIPP Panels:
First Quarter 1984 to Third Quarter 1991
Real 3rd quarter 1991 index

