

Appendix D. Description of the SIPP 1990 Panel File and Data Quality

DESCRIPTION OF SIPP 1990 PANEL FILE

The estimates presented in this report are based on the SIPP 1990 panel file. This file contains monthly data for persons over a 32-month period. The staggered SIPP design (described in appendix A) means that the actual reference periods are October 1989 to May 1992, November 1989 to June 1992, December 1989 to July 1992, and January 1990 to August 1992. The period covered by the 1990 longitudinal panel file consists of 32 interview months (eight interviews) for rotations 1, 2,

3, and 4. Data from all four rotation groups are available only for the reference period January 1990 through May 1992.

Each person in the panel file has been assigned three weights: a weight for calendar year 1990, a weight for calendar year 1991, and a weight for the 32-month reference period. In order to receive a non-zero weight, a person must have an observation for each month of the relevant reference period or have a complete set of observations up until the time he or she died, became institutionalized, or moved to Armed Forces barracks or out of the country. The data shown in this report are

Table D-1. Percent Distribution: Three Categories of Sample Persons: 1990 SIPP Panel

Characteristic	Complete set of interviews obtained ¹	Interviewed in first wave, left sample for reasons other than death, institutionalization, or a move to Armed Forces barracks or out of the country	Not a member of sample household during first wave; became members in second or later waves
Total	43,799 (100.0)	14,489 (100.0)	10,827 (100.0)
SEX			
Male	47.0	50.4	51.0
Female	53.0	49.6	49.0
AGE AT FIRST INTERVIEW			
Under 18 years	28.6	27.4	38.0
18 to 24 years	8.6	15.6	20.8
25 to 44 years	31.5	33.3	27.5
45 to 64 years	18.6	16.5	10.4
65 years and over	12.7	7.3	3.2
PROGRAM PARTICIPATION, FIRST MONTH IN SAMPLE			
Persons 18 years and over	31,267 (100.0)	10,523 (100.0)	6,711 (100.0)
Participated in major assistance program	9.1	10.5	11.4
AFDC or general assistance	2.0	2.9	2.7
Food stamps	4.4	5.1	5.8
Medicaid	5.5	5.9	7.1
Public/subsidized housing	3.9	4.8	2.8
SSI	2.3	1.8	2.2
Did not participate	90.9	89.5	88.6
Covered by private health insurance	79.7	70.2	67.9
Health insurance provided through employer	44.2	39.2	39.5
Not covered by private health insurance	20.4	29.8	32.1

¹Includes 1,832 persons who died, were institutionalized, or moved to Armed Forces barracks or out of the country during the 32-month period.

affected if characteristics of persons with an incomplete set of observations differed from those with a complete set.

Table D-1 shows three categories of sample persons by sex, age, and program participation status. The numbers in the table are unit counts; they are not weighted. The category "complete set of interviews obtained" includes 43,799 persons. The next category, "Interviewed in first wave, left sample for reasons other than death, institutionalization, or a move to Armed Forces barracks or out of the country" includes 14,489 persons. The final category includes 10,827 persons who were not members of a SIPP household during the first wave of interviews, but who subsequently became members of a sample household.

A comparison of the first two columns shows that the characteristics of those who completed the full set of interviews are reasonably close to the characteristics of those who dropped out of the sample. The major differences in the age distribution are for young adults and for the elderly. Young adults are underrepresented and the elderly are overrepresented in the group of persons who completed the full set of interviews. The data in table D-1 are, as noted, unweighted, and any potential problem caused by unrepresentative age distributions are minimized when the file is weighted to independent controls.

TIME-IN-SAMPLE BIAS

The use of the panel file to obtain estimates for 1990 and 1991 raises the issue of time-in-sample bias. There is ample evidence that certain measures vary according to the number of times the respondent has been visited. In the CPS, for example, the measured unemployment rate is always higher for the group of households being interviewed for the first time than for the groups being interviewed for the second or later times.

Time-in-sample bias arises when a person's response to a survey question (or the interviewer's method of asking a question) is influenced by what occurred in a previous visit. The overlapping SIPP sample design provides the data that allows for an examination of the presence of time-in-sample bias in SIPP estimates. That is, it is possible in SIPP to obtain estimates for a given time period from two or more separate panels and the amount of time respondents will have spent in the SIPP panel will differ for each of the panels. For example, estimates for each of the four quarters of 1991 can be obtained from both the 1990 and 1991 panels (respondents in the 1990 panel will have had more visits).

Quarterly estimates for the years 1984 to 1992 are shown in table D-2. Estimates from each panel file are shown separately for comparison. The estimates shown are the number of poor nonfarm households as well as the number of households receiving various government benefits.

The figures in table D-2 provide very little evidence regarding the existence of time-in-sample bias for several reasons. First, most of the observed differences are smaller than the differences that could be explained by sampling error. Second, a single observation is not sufficient to identify a pattern of bias. Third, differences may be attributable to attrition bias rather than time-in-sample bias. In spite of these qualifications, however, the observed relationships offer some reason to be cautious in interpreting the differences between the 1990 and 1991 estimates that have been presented earlier in this report.

OTHER ISSUES OF DATA QUALITY

Two major determinants of the quality of income data collected in household surveys are the magnitude of missing responses and the accuracy of the responses that are provided. This appendix has been included to supply information concerning nonresponse rates for selected income questions, the average amounts of income reported in the survey or assigned in the imputation of missing responses, and the extent to which the survey figures underestimate numbers of income recipients and amounts of income received.

Nonresponse in this discussion refers to missing responses to specific questions or "items" on the questionnaire. Noninterviews or complete failure to obtain cooperation from any household member have not been considered in this examination of nonresponse rates. Adjustments to account for noninterviews are made by proportionally increasing the survey weights of interviewed households. Missing responses to specific questions are assigned a value in the imputation phase of the data processing operation.

Nonresponse is a very important factor in assessing the quality of survey data. Nonresponses to income questions cannot be considered random since experience has shown that persons with the highest nonresponse rates have reported characteristics such as education levels and occupations that, in general, differ from population averages. The most frequent causes of nonresponse are the inability of the respondent to answer the question because of either a (1) lack of knowledge or (2) refusal to answer. The first reason is especially important in situations of proxy response when one household member answers questions for another household member not present at the time of the interview. The practice of accepting proxy interviews from household members deemed "qualified" to answer is a standard procedure in the CPS and most other surveys conducted by the Bureau. During the eight interviews of the 1990 panel, an average of 36 percent of the interviews were taken from proxy respondents.

Nonresponses are assigned values prior to producing estimates from the survey data. The procedure used to assign or impute responses for missing data for SIPP

Table D-2. Average Monthly Nonfarm Poverty and Household Participation in Means-Tested and Nonmeans-Tested Government Programs—SIPP Panels: First Quarter 1984 to Third Quarter 1992

[Numbers in thousands]

Panel, year, and quarter	All households	Non-farm poor households	Re-ceived government benefits	Re-ceived means-tested benefits	Re-ceived AFDC or other non-SSI PA ¹	Re-ceived SSI	Re-ceived food stamps	Covered by Medicaid	Lived in public or subsidized housing	Re-ceived non-means-tested benefits	Re-ceived Social Security or RR ²	Re-ceived unemployment compensation
1984 Panel												
1984: 1st quarter.....	83,643	11,722	39,050	16,052	3,870	2,880	6,462	7,593	3,615	30,578	23,479	2,659
2nd quarter.....	84,002	11,124	37,712	14,460	3,832	2,975	6,303	7,559	3,669	29,973	23,525	2,047
3rd quarter.....	84,609	11,050	36,674	13,188	3,565	3,002	5,989	7,242	3,670	29,747	23,475	1,784
4th quarter.....	84,945	11,160	38,347	15,276	3,585	3,008	6,107	7,211	3,584	30,308	23,559	2,240
1985: 1st quarter.....	84,948	10,922	39,158	15,403	3,763	3,020	6,230	7,458	3,546	31,238	23,821	2,944
2nd quarter.....	85,614	10,783	37,725	13,921	3,651	3,085	5,955	7,264	3,486	30,485	23,955	2,062
3rd quarter.....	86,249	10,872	37,143	13,013	3,598	3,036	5,886	7,370	3,502	30,339	23,938	1,757
4th quarter.....	86,585	10,688	38,547	15,274	3,622	3,064	5,839	7,461	3,583	30,379	23,864	1,824
1985 Panel												
1985: 1st quarter.....	85,228	11,585	(NA)	(NA)	3,536	2,990	5,999	7,277	3,751	(NA)	23,559	2,872
2nd quarter.....	85,875	10,929	(NA)	(NA)	3,554	3,116	5,808	7,460	3,778	(NA)	23,781	2,232
3rd quarter.....	86,561	11,088	(NA)	(NA)	3,526	3,211	5,624	7,501	3,844	(NA)	23,838	1,883
4th quarter.....	86,832	10,978	(NA)	(NA)	3,499	3,200	5,676	7,517	3,889	(NA)	23,929	2,075
1986: 1st quarter.....	86,911	10,890	(NA)	(NA)	3,591	3,114	5,800	7,589	3,868	(NA)	24,145	2,617
2nd quarter.....	86,940	10,463	(NA)	(NA)	3,502	3,168	5,617	7,456	3,847	(NA)	24,130	2,112
3rd quarter.....	87,180	10,873	(NA)	(NA)	3,462	3,201	5,573	7,499	3,813	(NA)	24,150	2,037
4th quarter.....	87,607	10,872	(NA)	(NA)	3,516	3,181	5,641	7,575	3,710	(NA)	24,264	2,206
1986 Panel												
1986: 1st quarter.....	87,134	11,711	40,711	16,422	3,753	3,107	6,588	7,593	4,172	31,702	24,013	2,443
2nd quarter.....	87,192	11,185	40,125	16,034	3,858	3,270	6,524	7,901	4,347	31,272	24,102	2,001
3rd quarter.....	87,313	11,144	39,427	14,999	3,680	3,290	6,180	7,879	4,239	31,448	24,310	2,062
4th quarter.....	87,721	11,081	40,620	16,418	3,780	3,270	6,212	8,041	4,199	31,697	24,523	1,985
1987: 1st quarter.....	88,118	11,471	41,243	16,654	3,885	3,347	6,340	8,265	4,172	32,268	24,567	2,356
2nd quarter.....	88,367	11,022	39,781	15,123	3,848	3,395	6,218	8,278	4,179	31,619	24,612	1,720
3rd quarter.....	88,645	10,839	39,358	14,975	3,794	3,482	5,900	8,288	4,204	31,067	24,572	1,442
4th quarter.....	89,058	10,634	40,846	16,431	3,718	3,399	5,789	8,189	4,160	31,532	24,747	1,579
1987 Panel												
1987: 1st quarter.....	88,131	11,291	40,543	15,793	3,717	3,075	6,230	7,707	3,883	32,216	24,529	2,616
2nd quarter.....	88,237	10,691	39,345	14,857	3,597	3,222	6,142	7,795	3,979	31,470	24,702	1,869
3rd quarter.....	88,601	10,648	38,681	13,605	3,485	3,208	5,881	7,759	3,861	31,670	25,198	1,624
4th quarter.....	89,067	10,558	40,395	15,748	3,425	3,194	5,833	7,818	3,941	32,073	25,150	2,179
1988: 1st quarter.....	89,719	10,630	40,853	15,921	3,607	3,172	5,916	7,919	4,053	32,282	25,410	1,996
2nd quarter.....	89,963	10,212	39,418	14,527	3,534	3,159	5,824	7,986	4,006	31,615	25,403	1,448
3rd quarter.....	90,401	10,354	38,201	13,285	3,382	3,150	5,706	7,909	4,001	31,279	24,947	1,296
4th quarter.....	90,481	10,461	40,102	15,314	3,329	3,186	5,813	7,876	4,012	31,952	25,228	1,495
1988 Panel												
1988: 1st quarter.....	89,858	11,547	41,438	16,741	3,318	3,496	5,996	7,882	4,488	32,334	25,399	2,164
2nd quarter.....	90,066	11,020	40,078	15,160	3,356	3,747	5,948	8,265	4,557	32,012	25,456	1,655
3rd quarter.....	90,488	11,154	39,634	14,320	3,375	3,793	5,851	8,530	4,515	32,262	25,641	1,593
4th quarter.....	90,517	10,855	41,316	16,860	3,417	3,790	6,038	8,556	4,563	32,297	25,573	1,540
1989: 1st quarter.....	91,144	10,677	41,775	16,829	3,477	3,682	6,133	8,405	4,746	32,796	25,762	2,087
2nd quarter.....	91,614	10,548	40,740	15,272	3,362	3,624	5,892	8,274	4,716	32,753	25,959	1,565
3rd quarter.....	91,841	10,858	39,866	14,142	3,275	3,645	5,856	8,455	4,556	32,549	25,975	1,380
4th quarter.....	92,103	11,144	42,048	16,380	3,265	3,721	6,108	8,730	4,252	32,579	26,369	1,382

See footnotes at the end of the table.

Table D-2. Average Monthly Nonfarm Poverty and Household Participation in Means-Tested and Nonmeans-Tested Government Programs—SIPP Panels: First Quarter 1984 to Third Quarter 1992—Continued

[Numbers in thousands]

Panel, year, and quarter	All households	Non-farm poor households	Received government benefits	Received means-tested benefits	Received AFDC or other non-SSI PA ¹	Received SSI	Received food stamps	Covered by Medicaid	Lived in public or subsidized housing	Received non-means-tested benefits	Received Social Security or RR ²	Received unemployment compensation
1990 Panel												
1990: 1st quarter.....	91,776	10,864	42,687	16,690	3,523	3,239	5,939	8,348	4,576	33,472	26,257	2,407
2nd quarter.....	92,308	10,462	41,827	15,303	3,696	3,315	6,184	8,607	4,470	33,397	26,412	2,011
3rd quarter.....	92,722	11,055	41,419	14,397	3,711	3,354	6,254	8,871	4,452	33,686	26,745	1,921
4th quarter.....	92,421	11,381	43,554	17,130	3,722	3,369	6,334	9,014	4,431	33,939	26,529	2,288
1991: 1st quarter.....	93,058	11,548	44,708	17,776	3,879	3,482	6,631	9,326	4,502	34,922	26,702	3,354
2nd quarter.....	93,278	11,367	43,316	16,230	4,004	3,418	6,664	9,462	4,490	34,434	26,732	2,818
3rd quarter.....	93,443	11,555	42,568	15,212	4,150	3,475	6,750	9,853	4,513	34,364	26,705	2,681
4th quarter.....	93,289	11,545	44,975	17,921	4,045	3,516	6,857	10,015	4,452	34,944	26,642	3,183
1992: 1st quarter.....	93,879	11,679	46,292	18,448	4,089	3,744	7,152	10,339	4,504	36,286	26,943	4,244
1991 Panel												
1991: 1st quarter.....	93,313	12,453	44,115	18,203	4,051	3,593	6,814	9,652	4,836	34,047	26,189	3,120
2nd quarter.....	93,545	11,876	43,439	17,131	4,152	3,799	6,945	9,996	4,845	34,103	26,299	2,861
3rd quarter.....	93,774	12,162	42,903	16,149	4,003	3,936	6,975	10,313	4,830	34,214	26,369	2,714
4th quarter.....	93,649	12,294	45,301	18,677	3,927	3,902	7,055	10,281	4,706	34,846	26,456	3,236
1992: 1st quarter.....	94,157	12,422	46,455	18,962	4,057	3,957	7,358	10,533	4,871	36,180	26,685	4,502
2nd quarter.....	94,588	12,240	44,958	17,557	4,062	4,173	7,248	10,678	4,772	35,509	26,591	4,076
3rd quarter.....	94,095	12,291	43,665	16,556	3,987	4,152	7,199	10,813	4,621	34,902	26,533	3,550

NA Not available.

¹AFDC = Aid to Families with Dependent Children; SSI = Supplemental Security Income; PA = Public Assistance.

²RR = Railroad Retirement.

is commonly referred to as the "hot deck" imputation method. This process assigns values reported in the survey by respondents to nonrespondents. The respondent from whom the value is taken is termed the "donor." Values from donors are stored in a matrix defined by demographic and economic data available for both donors and nonrespondents. Each cell of the matrix defines a unique combination of demographic and economic characteristics. For example, the imputation of an amount for monthly wage and salary income is based on eight different variables: (1) occupation, (2) sex, (3) age, (4) race, (5) educational attainment, (6) weeks worked, (7) usual hours worked per week, and (8) place of residence.

The second important determinant of data quality and probably the one examined most closely by users of the income data collected in household surveys is the accuracy of reported (and imputed) amounts. In general, household surveys have a tendency to underestimate the number of persons receiving income and the

average amount received. These problems result for a variety of reasons including random response error, misreporting of sources of income, failure to report the receipt of income from a specified source, and failure to report the full amount received. The net effect of these kinds of problems is, for most income types, underestimation or underreporting of income amounts. The extent of underreporting is measured by comparing survey estimates with independently derived estimates, usually based on administrative data that are generally more reliable than the estimates derived from the survey. It should be noted that the independent estimates are subject to errors themselves. In addition, independent estimates do not reflect income attributable to the "underground" economy, some of which may be reported in the survey. AFDC and food stamp participation for January 1992 were 25.3 and 23.4 percent lower, respectively, in the 1990 SIPP panel than the numbers from the administrative records of the Departments of Health and Human Services and Agriculture.