

Appendix D. 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 second and third interview periods of SIPP, about 34 percent of the interviews were taken from proxy respondents.

The magnitude of nonresponse is generally presented in terms of a nonresponse rate computed by dividing the number of nonresponses by the total number of responses that should have been provided. The first two columns of table D-1 show the number of persons with income and nonresponse rate for a selected group of income amount questions from SIPP for the first quarter of 1984. Nonresponse rates for the March 1984 CPS based on annual income amount questions are shown in the third column.

Table D-1. Persons Nonresponse Rates for SIPP and the March 1984 CPS, and Median Monthly Amounts Reported and Imputed, for Selected Income Types

Income type	SIPP 1984 first quarter monthly average		March 1984 CPS nonresponse rate for amounts received	SIPP median monthly average amounts	
	Number with income (thous.)	Nonresponse rate for amounts received		Reported	Imputed
Wage or salary.....	94,409	7.2	17.4	\$1,127	\$1,197
Self-employment income.....	8,014	16.8	25.2	794	1,110
Federal Supplemental Security Income.....	3,362	7.6	16.5	209	244
Social Security income.....	32,370	10.8	20.1	397	404
Aid to Families with Dependent Children.....	3,129	6.1	13.4	293	241
Unemployment compensation.....	2,982	10.1	19.0	401	341
Company or union pensions.....	7,807	13.9	22.6	239	266
Food stamp allotment.....	6,917	5.2	12.7	101	101
Veterans' compensation or pensions.....	3,546	11.3	16.6	126	119

The SIPP nonresponse rates ranged from a low of about 5 percent for food stamps to about 17 percent for self-employment income. These rates were computed by dividing the number of persons with missing responses on the amount received by the total number with either a missing or reported amount for that income type.

The SIPP nonresponse rates for first-quarter average monthly amounts contrast sharply with the higher nonresponse rates encountered in the March CPS. The rates for the CPS ranged from a low of 13 percent for food stamp allotments and Aid to Families with Dependent Children to 25 percent for self-employment income. The major emphasis given to complete and accurate income information in SIPP and 4-month recall period are two factors that have contributed to the significantly lower nonresponse rates in the SIPP.

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 are of a type commonly referred to as a "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. These were 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 last two columns in table D-1 compare median reported and imputed income amounts for SIPP monthly averages, first quarter 1984. In most cases the differences between imputed amounts are not statistically significant. The differences between reported and imputed median wage and salary amounts (\$1,127 and \$1,197, respectively) and reported and imputed

median self-employment income (\$794 and \$1,110, respectively) and reported and imputed median Aid to Families with Dependent Children amounts (\$293 and \$241, respectively) were statistically significant.

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.

Table D-2 contains comparisons of SIPP estimates of the number of persons receiving specific income sources with independent estimates derived from various administrative sources. Table D-3 shows similar comparisons based not on the number of recipients but on the aggregate amount of income received. Data in both of these tables are preliminary and subject to revision.

The comparisons in table D-2 are limited to some of the major transfer programs for which administrative data are available for the January-March 1984 period. Adjustment factors were applied to these administrative figures in order to arrive at the independent estimates for the SIPP noninstitutional population eligible for interview. The adjustment factors used were based on procedures developed by Mathematical

Table D-2. Comparison of Estimated Number of Income Recipients, for Selected Income Types, First Quarter 1984: SIPP vs. Independently Derived Estimates

(Numbers in thousands)

Income type	Monthly average recipients		SIPP as a percent of independent estimate
	SIPP estimate	Independent estimate	
Federal Supplemental Security Income.....	3,362	3,545	94.8
Social Security income.....	32,370	33,254	97.3
Aid to Families with Dependent Children ¹	3,129	3,698	84.6
Unemployment compensation.....	2,982	3,608	82.6
Food stamp allotment.....	19,448	21,407	90.8
Veteran's compensation or pensions ¹	3,546	3,913	90.6

¹Excludes dependents covered by payments.

Table D-3. Comparison of Estimated Aggregate Income Amounts Received, for Selected Income Types, First Quarter 1984: SIPP vs. Independently Derived Estimates

(Monthly averages. Figures in millions of dollars)

Income type	SIPP estimate	Independent estimate	SIPP as a percent of the independent estimate
Wage or salary.....	134,120	143,940	93.2
Self-employment income.....	16,412	(NA)	(X)
Federal Supplemental Security income.....	744	772	96.4
Social Security income.....	13,113	13,054	100.5
Aid to Families with Dependent Children.....	989	1,159	85.3
Unemployment compensation.....	1,206	1,415	85.2
Company or union pensions.....	2,861	(NA)	(X)
Food stamp allotment.....	782	918	85.2
Veterans' compensation or pensions.....	814	1,044	78.0

NA Not available.

X Not applicable.

Policy Research, Inc., for deriving independent estimates for the 1979 ISDP research panel.

Survey underestimates of income recipients ranged from about 17 percent for State unemployment compensation payments and 15 percent for Aid to Families with Dependent Children to about 3 percent for Social Security recipients.

The underreporting for Aid to Families with Dependent Children is related to misclassification of this income type as other types of public assistance or welfare. A total of 1,084,000 persons reported receiving general assistance and 178,000 reported receiving other types of welfare payments for the first quarter. A significant number of these cases are actually payments from the Aid to Families with Dependent Children program. This particular problem was also encountered and documented in the developmental ISDP.

Table D-3 provides comparisons of SIPP and independent estimates of the aggregate amount of income received for the total noninstitutional population for the first quarter of 1984. Nonseasonally adjusted, monthly independent estimates for wage and salary income is not available. The estimate shown for wage and salary is based on Bureau of Economic Analysis seasonally adjusted, annual rate estimates for the first quarter divided by 12. Other independent

estimates shown in tables D-2 and D-3 are based on various sources including the Social Security Bulletin and unpublished figures from the Department of Health and Human Services, the Department of Agriculture, and the Veterans' Administration.

In most cases the comparisons in table D-3 on aggregate amounts for the first quarter parallel the figures in table D-2 for estimated number of recipients. The comparison for wage and salary income is difficult to interpret because the independent estimate is seasonally adjusted. A monthly independent estimate for self-employment income is not available because the self-employment income estimates are based on different concepts. The SIPP figure is based on the "salary" and other income received from the business by the owners. More refined comparisons between SIPP estimates and estimates derived from independent sources will be made in future reports.

Table D-4 shows the monthly averages for the number of income recipients and aggregate amounts of income received for the first quarter 1984 for the total population and the non-farm population. Most of the largest sources of income have been included in this table.

Table D-4. Persons 15 Years Old and Over Receiving Income and Aggregate Amount Received, for Selected Income Types: Monthly Average, First Quarter 1984

Income type	Total		Nonfarm	
	Number with income (thousands)	Aggregate amount (millions of dollars)	Number with income (thousands)	Aggregate amount (millions of dollars)
Wage or salary.....	94,409	134,120	92,829	132,397
Self-employment income.....	8,014	16,412	7,398	13,250
Social Security income.....	32,370	13,113	31,418	12,760
Federal Supplemental Security Income..	3,362	744	3,322	740
State unemployment compensation.....	2,982	1,206	2,934	1,191
Veterans' compensation or pensions....	3,546	814	3,477	807
Workers' compensation.....	657	464	639	457
Aid to Families with Dependent Children.....	3,129	989	3,102	983
General assistance.....	1,084	212	1,083	212
Child support.....	3,085	757	3,064	754
Alimony.....	519	208	519	208
Company or union pensions.....	7,807	2,861	7,712	2,825
Federal employee pensions.....	1,822	1,681	1,799	1,660
U.S. military pensions.....	1,275	1,293	1,264	1,281
State government employee pensions....	1,956	1,020	1,886	984
Local government employee pensions....	825	433	812	423
Estates and trusts.....	272	372	262	349
Income from paid-up life insurance or other annuities.....	782	147	768	145
Money from relatives or friends.....	1,378	520	1,371	511
Interest income from regular savings accounts, money market deposit accounts, certificates of deposit or other savings certificates, and interest-bearing checking accounts...	100,867	6,934	98,064	6,711
Interest income from money market funds, U.S. government securities, municipal or corporate bonds, and other interest-bearing assets.....	10,244	1,995	10,047	1,935
Interest on mortgages.....	3,371	925	3,257	849
Dividends.....	20,780	3,095	20,314	2,967
Net rental income.....	9,860	1,427	9,392	1,302
Income from royalties and other financial investments.....	2,884	1,612	2,709	1,550