

## 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 occupa-

tions 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 fourth and part of the third and fifth interview periods of SIPP, about 36 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 third 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 third 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.....	97,789	7.5	17.4	\$1,144	\$1,247
Self-employment income.....	8,489	16.0	25.2	849	1,141
Federal Supplemental Security income.....	3,549	8.1	16.5	203	235
Social Security income.....	32,376	11.7	20.1	402	407
Aid to Families with Dependent Children.....	2,973	6.5	13.4	290	243
Unemployment compensation.....	1,927	10.4	19.0	368	314
Company or union pensions.....	8,047	12.8	22.6	241	247
Food stamp allotment.....	6,416	6.7	12.7	98	78
Veterans' compensation or pensions.....	3,435	11.9	16.6	128	100

The SIPP nonresponse rates ranged from a low of about 7 percent for food stamps to about 16 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 third-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, third quarter 1984. The differences between reported and imputed median amounts were statistically significant at the 95-percent confidence level for wage and salary income, self-

employment income, Aid to Families with Dependent Children, food stamp allotments, and veterans' compensation or pensions.

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 July-September 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 Numbers of Income Recipients for Selected Income Types, Third 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.....	3,549	3,612	98.3
Social Security income.....	32,376	33,197	97.5
Aid to Families with Dependent Children <sup>1</sup> .....	2,973	3,624	82.0
Unemployment compensation.....	1,927	2,456	78.5
Food stamp allotment.....	17,921	19,857	90.3
Veteran's compensation or pensions <sup>1</sup> .....	3,435	3,824	89.8

<sup>1</sup>Excludes dependents covered by payments.

**Table D-3. Comparison of Estimated Aggregate Income Amounts Received, for Selected Income Types, Third 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.....	141,914	149,121	95.2
Self-employment income.....	15,999	(NA)	(X)
Federal Supplemental Security income.....	774	785	98.6
Social Security income.....	13,299	13,129	101.3
Aid to Families with Dependent Children.....	949	1,183	80.2
Unemployment compensation.....	762	949	80.3
Food stamp allotment.....	711	840	84.6
Veterans' compensation or pensions.....	798	1,046	76.3

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 22 percent for State unemployment compensation payments and 18 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 967,000 persons reported receiving general assistance and 180,000 reported receiving other types of welfare payments for the third 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 third 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 third 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 third 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 third 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, Third 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.....	97,789	141,914	95,990	139,794
Self-employment income.....	8,489	15,999	7,888	14,047
Social Security income.....	32,376	13,299	31,423	12,938
Federal Supplemental Security income..	3,549	774	3,512	770
State unemployment compensation.....	1,927	762	1,902	753
Veterans' compensation or pensions....	3,435	798	3,364	788
Workers' compensation.....	628	457	619	447
Aid to Families with Dependent Children.....	2,973	949	2,948	944
General assistance.....	967	186	967	186
Child support.....	3,115	751	3,091	748
Alimony.....	432	233	432	233
Company or union pensions.....	8,047	2,901	7,926	2,857
Federal employee pensions.....	1,811	1,655	1,780	1,630
U.S. military pensions.....	1,323	1,316	1,317	1,311
State government employee pensions....	2,122	1,105	2,053	1,069
Local government employee pensions....	860	464	845	453
Estates and trusts.....	287	304	278	301
Income from paid-up life insurance or other annuities.....	699	156	685	153
Money from relatives or friends.....	1,260	531	1,251	530
Interest income from regular savings accounts, money market deposit accounts, certificates of deposit or other savings certificates, and interest-bearing checking accounts...	102,271	6,544	99,522	6,311
Interest income from money market funds, U.S. government securities, municipal or corporate bonds, and other interest-bearing assets.....	10,596	1,579	10,372	1,531
Interest on mortgages.....	3,411	1,034	3,315	993
Dividends.....	20,148	3,458	19,714	3,385
Net rental income.....	9,841	1,374	9,375	1,307
Income from royalties and other financial investments.....	2,725	2,746	2,562	2,692