Appendix C. Source and Accuracy of Estimates

SOURCE OF DATA

The SIPP universe is the noninstitutionalized resident population living in the United States. This population includes persons living in group quarters, such as dormitories, rooming houses, and religious group dwellings. Crew members of merchant vessels, Armed Forces personnel living in military barracks, and institutionalized persons, such as correctional facility inmates and nursing home residents, are not eligible to be in the survey. Also not eligible are United States citizens residing abroad. Foreign visitors who work or attend school in this country and their families are eligible; all others are not eligible. With the exceptions noted above, field representatives interview eligible persons who are at least 15 years of age at the time of the interview.

The 1991 SIPP panel is located in 230 Primary Sampling Units (PSU's) each consisting of a county or a group of contiguous counties. Within these PSU's, we systematically selected expected clusters of two living quarters from lists of addresses prepared for the 1980 decennial census to form the bulk of the sample. To account for living quarters built within each of the sample areas after the 1980 census, we selected a sample containing clusters of four living quarters from permits issued for construction of residential living quarters up until shortly before the beginning of the panel.

In jurisdictions that have incomplete addresses or do not issue building permits, we sampled small land areas, listed expected clusters of four living quarters, and then subsampled. In addition, we selected a sample of living quarters from a supplemental frame that included living quarters identified as missed in the 1980 census.

The first interview occurred during February, March, April, and May of 1991. Interviews for approximately one-fourth of the sample took place in each of these months. For the remainder of the panel, interviews for each person occurred every 4 months. At each interview the reference period was the 4 months preceding the interview month.

Occupants of about 93 percent of all eligible living quarters participated in the first interview of the panel. For later interviews, field representatives interviewed only original sample persons (those in Wave 1 sample households and interviewed in Wave 1) and persons living with them. The Bureau automatically designated all first wave noninterviewed households as noninterviews for all subsequent interviews. We followed original sample persons if they moved to a new address, unless the new address was more than 100 miles from a SIPP sample area. If the original sample persons moved farther than 100 miles from a SIPP sample area, we attempted telephone interviews. When the original sample persons moved to remote parts of the country and were unreachable by telephone, moved without leaving a forwarding address, or refused the interview, additional noninterviews resulted.

As a part of most waves, we cover subjects that are important to meet SIPP goals and do not require repeated measurement during the panel. The data on these subjects are of particular interest to data users and policy makers. We cover these subjects once during the panel or annually. By collecting data once for the panel or annually, we reduce respondent burden. We call a specific set of questions on a subject a topical module. For this report, the topical module analyzed includes questions on the living arrangements of children. We implemented them in wave 2 of the 1991 panel.

Noninterviews. Tabulations in this report were drawn from interviews conducted from June through September 1991. Table 1 summarizes information on nonresponse for the interview months in which we collected the data used to produce this report.

Table 1. **Household Sample Size by Month and Interview Status**

Month	Eligible	Interviewed	Noninterviewed	Nonresponse rate (percent)
June 1991	4,000	3,400	600	15.0
July 1991	3,900	3,400	500	13.6
August 1991	4,000	3,400	500	13.0
September 1991	3,900	3,400	500	13.7

[1]Because of rounding of all numbers to the nearest 100, there are some inconsistencies. We calculated the percentage using unrounded numbers.

Some respondents do not respond to some of the questions. Therefore, the overall nonresponse rate for some items is higher than the nonresponse rates in table 1. For more discussion of nonresponse, see the *Quality Profile for the Survey of Income and Program Participation,* May 1990, by T. Jabine, K. King, and R. Petroni, available from Customer Services, Data Users Services Division, of the U.S. Census Bureau (301-763-6100).

WEIGHTING PROCEDURE

We derived SIPP person weights in each panel from several stages of weight adjustments. In the first wave, we gave each person a base weight equal to the inverse of his/her probability of selection. For each subsequent interview, the Bureau gave each person a base weight that accounted for following movers.

We applied a factor to each interviewed person's weight to account for the SIPP sample areas not having the same population distribution as the strata they are from.

We applied a noninterview adjustment factor to the weight of every occupant of interviewed households to account for persons in noninterviewed occupied households that were eligible for the sample. (The Bureau treated individual nonresponse within partially interviewed households with imputation. We made no special adjustment for noninterviews in group quarters.)

The Bureau used complex techniques to adjust the weights for nonresponse. For a further explanation of the techniques used, see the *Nonresponse Adjustment Methods for Demographic Surveys at the U.S. Bureau of the Census*, November 1988, Working paper 8823, by R. Singh and R. Petroni. The success of these techniques in avoiding bias is unknown. An example of successfully avoiding bias is in "Current Nonresponse Research for the Survey of Income and Participation" (paper by Petroni, presented at the Second International Workshop on Household Survey Nonresponse, October 1991).

We performed an additional stage of adjustment to persons' weights to reduce the mean square errors of the survey estimates. We accomplished this by ratio adjusting the sample estimates to agree with monthly Current Population Survey (CPS) type estimates of the civilian (and some military) noninstitutional population of the United States at the national level by demographic characteristics including age, sex, and race as of the specified date. The Bureau brought CPS estimates by age, sex, and race into agreement with adjusted estimates from the 1980 decennial census. Adjustments to the 1980 decennial census estimates reflect births, deaths, immigration, emigration, and changes in the Armed Forces since 1980. In addition, we controlled SIPP estimates to independent Hispanic controls and made an adjustment to assign equal weights to husbands and wives within the same household. We implemented all of the above adjustments for each reference month and the interview month.

ACCURACY OF ESTIMATES

We base SIPP estimates on a sample. The estimates may differ somewhat from the values obtained from administering a complete census using the same questionnaire, instructions, and enumerators. The difference occurs because with an estimate based on a sample survey two types of errors are possible: nonsampling and sampling. We can provide estimates of the magnitude of the SIPP sampling error, but this is not true of nonsampling error. The next few sections describe SIPP nonsampling error sources, followed by a discussion of sampling error, its estimation, and its use in data analysis.

Nonsampling variability. We attribute nonsampling errors to many sources, they include the following:

a. Inability to obtain information about all cases in the sample.

b. Definitional difficulties.

c. Differences in the interpretation of questions.

d. Inability or unwillingness on the part of the respondents to provide correct information.

e. Inability to recall information.

f. Errors made in collection (e.g., recording or coding the data).

g. Errors made in processing the data.

h. Errors made in estimating values for missing data.

i. Biases resulting from the differing recall periods caused by the interviewing pattern used.

j. Undercoverage.

We used quality control and edit procedures to reduce errors made by respondents, coders, and interviewers. More detailed discussions of the existence and control of nonsampling errors in the SIPP are in the *SIPP Quality Profile*.

Undercoverage in SIPP resulted from missed living quarters and missed persons within sample households. It is known that undercoverage varies with age, race, and sex. Generally, undercoverage is larger for males than for females and larger for Blacks than for non-Blacks. Ratio estimation to independent age-race-sex population controls partially corrects for the bias resulting from survey undercoverage. However, biases exist in the estimates when persons in missed households or missed persons in interviewed households have characteristics different from those of interviewed persons in the same age-race-sex group. Further, we did not adjust the independent population controls for undercoverage in the census.

A common measure of survey coverage is the coverage ratio, the estimated population before ratio adjustment divided by the independent population control. Table 2 shows CPS coverage ratios for age-sex-race groups for 1992. The CPS coverage ratios can exhibit some variability from month to month, but these are a typical set of coverage ratios. Other Census Bureau household surveys like the SIPP experience similar coverage.

Table 2. **1992 CPS Coverage Ratios**

Age	All persons			Non-Black		Black	
	Total	Males	Females	Males	Females	Males	Females
0-14	0.958	0.957	0.959	0.963	0.965	0.927	0.926
15	0.948	0.952	0.944	0.962	0.949	0.899	0.919
16	0.947	0.962	0.932	0.969	0.936	0.923	0.907
17	0.966	0.975	0.957	0.981	0.975	0.945	0.862
18	0.922	0.930	0.913	0.939	0.926	0.883	0.846
19	0.853	0.844	0.861	0.860	0.872	0.754	0.801
20-24	0.901	0.889	0.913	0.913	0.927	0.734	0.832
25-26	0.914	0.897	0.931	0.927	0.940	0.688	0.877
27-29	0.914	0.885	0.941	0.910	0.954	0.707	0.864
30-34	0.905	0.870	0.939	0.893	0.948	0.691	0.883
35-39	0.919	0.895	0.942	0.910	0.949	0.763	0.899
40-44	0.933	0.919	0.946	0.929	0.951	0.824	0.906
45-49	0.958	0.951	0.965	0.956	0.966	0.903	0.956
50-54	0.940	0.927	0.951	0.940	0.961	0.807	0.877
55-59	0.930	0.932	0.928	0.944	0.941	0.826	0.825
60-62	0.946	0.948	0.944	0.965	0.956	0.792	0.850
63-64	0.894	0.884	0.903	0.905	0.907	0.669	0.872
65-67	0.947	0.921	0.969	0.935	0.979	0.783	0.875
68-69	0.923	0.913	0.931	0.925	0.942	0.789	0.831
70-74	0.962	0.920	0.995	0.926	0.993	0.856	1.014
75-99	0.975	0.961	0.983	0.977	0.989	0.764	0.912
15+	0.929	0.912	0.944	0.928	0.953	0.782	0.883
0+	0.935	0.923	0.947	0.936	0.955	0.827	0.895

Comparability with other estimates. Exercise caution when comparing data from this report with data from other SIPP publications or with data from other surveys. Comparability problems are from varying seasonal patterns for many characteristics, different nonsampling errors, and different concepts and procedures. Refer to the *SIPP Quality Profile* for known differences with data from other sources and further discussion.

Sampling variability. Standard errors indicate the magnitude of the sampling error. They also partially measure the effect of some nonsampling errors in response and enumeration, but do not measure any systematic biases in the data. The standard errors mostly measure the variations that occurred by chance because we surveyed a sample rather than the entire population.

USES AND COMPUTATION OF STANDARD ERRORS

Confidence intervals. The sample estimate and its standard error enable one to construct confidence intervals, ranges that would include the average result of all possible samples with a known probability. For example, if we selected all possible samples and surveyed each of these under essentially the same conditions and with the same sample design, and if we calculated an estimate and its standard error from each sample, then—

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.645 standard errors below the estimate to 1.645 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from 1.960 standard errors below the estimate to 1.960 standard errors above the estimate would include the average result of all possible samples.

The average estimate derived from all possible samples is or is not contained in any particular computed interval. However, for a particular sample, one can say with a specified confidence that the confidence interval includes the average estimate derived from all possible samples.

Hypothesis testing. One may also use standard errors for hypothesis testing. Hypothesis testing is a procedure for distinguishing between population characteristics using sample estimates. The most common type of hypothesis tested is (1) the population characteristics are identical versus (2) they are different. One can perform tests at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Unless noted otherwise, all statements of comparison in the report passed a hypothesis test at the 0.10 level of significance or better. This means that, for differences cited in the report, the estimated absolute difference between parameters is greater than 1.645 times the standard error of the difference.

To perform the most common test, compute the difference X_A - X_B, where X_A and X_B are sample estimates of the characteristics of interest. A later section explains how to derive an estimate of the standard error of the difference X_A - X_B. Let that standard error be s_{DIFF}. If X_A - X_B is between -1.645 times s_{DIFF} and +1.645 times s_{DIFF}, no conclusion about the characteristics is justified at the 10-percent significance level. If, on the other hand, X_A - X_B is smaller than -1.645 times s_{DIFF} or larger than +1.645 s_{DIFF}, the observed difference is significant at the 10-percent level. In this event, it is commonly accepted practice to say that the characteristics are different. Of course, sometimes this conclusion will be wrong. When the characteristics are, in fact, the same, there is a 10-percent chance of concluding that they are different.

Note that as we perform more tests, more erroneous significant differences will occur. For example, at the 10-percent significance level, if we perform 100 independent hypothesis tests in which there are no real differences, it is likely that about 10 erroneous differences will occur. Therefore, interpret the significance of any single test cautiously.

Note concerning small estimates and small differences. We show summary measures in the report only when the base is 200,000 or greater. Because of the

large standard errors involved, there is little chance that the estimates will reveal useful information when computed on a base smaller than 200,000. Also, nonsampling error in one or more of the small number of cases providing the estimate can cause large relative error in that particular estimate. We show estimated numbers, however, even though the relative standard errors of these numbers are larger than those for the corresponding percentages. We provide smaller estimates primarily to permit such combinations of the categories as serve each user's needs. Therefore, be careful in the interpretation of small differences since even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test.

Standard error parameters and tables and their use. Most SIPP estimates have greater standard errors than those obtained through a simple random sample because we sampled clusters of living quarters for the SIPP. To derive standard errors at a moderate cost and applicable to a wide variety of estimates, we made a number of approximations. We grouped estimates with similar standard error behavior and developed two parameters (denoted "a" and "b") to approximate the standard error behavior of each group of estimates. Because the actual standard error behavior was not identical for all estimates within a group, the standard errors we computed from these parameters provide an indication of the order of magnitude of the standard error for any specific estimate. The "a" and "b" parameters found in table 3 are for Wave 2 1991 panel estimates of the living arrangements of children.

For those users who wish further simplification, we also provide general standard errors in tables 4 and 5. The standard errors taken from these tables are less accurate. Methods for using the parameters and tables for computation of standard errors are given in the following sections.

Standard errors of estimated numbers. There are two ways to compute the approximate standard error, s_x, of an estimated number shown in this report. The first method is to look up the standard error of the estimate obtained by interpolation from table 4. Alternatively, approximate s_x using the formula

$$s_x = \sqrt{ax^2 + bx} \qquad (1)$$

from which we calculated the standard errors in table 4. Here x is the size of the estimate and "a" and "b" are the parameters in table 3. Direct calculation with formula 1 will provide more accurate results than interpolating values from table 4.

Illustration. Suppose SIPP estimates show that 13,955,000 children under age 18 lived in single-mother families in 1991. The "a" and "b" parameters from table 3 and the appropriate general standard error found by interpolation from table 4 are a = -0.0001340, b = 7,514, s_x = 279,000.

The 90-percent confidence interval as shown is from 13,496,000 to 14,414,000. Using formula (1), the approximate standard error is

$$\sqrt{(-0.0001340)(13{,}955{,}000)^2 + (7{,}514)(13{,}955{,}000)} = 281{,}000$$

The 90-percent confidence interval as shown is from 13,493,000 to 14,417,000. Therefore, a conclusion that the average estimate derived from all possible samples lies within a range computed in this way would be correct for roughly 90 percent of all samples.

Standard errors of estimates percentages. The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends on the size of the percentage and its base. One way to find the approximate standard error, $s_{(x,p)}$, of an estimated percentage p, which is the percentage of persons with a particular characteristic, is by interpolating values from table 5.

Alternatively, approximate it by the formula—

$$s_{(x,p)} = \sqrt{\frac{b}{x}(p)(100-p)} \qquad (2)$$

from which we calculated the standard errors in table 5. Here x is the total number of persons in the base of the percentage, p is the percentage ($0 \leq p \leq 100$), and b is the "b" parameter in table 3. Use of this formula will give more accurate results than interpolating values from table 5.

Illustration. Suppose that SIPP estimates 1 out of every 5 children or 20 percent, in single-mother families lived with at least one adult male in the household. The base for this percentage is 13,955,000. The "b" parameter from table 3 and the appropriate general standard error found by interpolation from table 5 are

$$b = 7{,}514 \text{ and } s_x = 0.9 \text{ percent}$$

Using formula (2), the approximate standard error is

$$\sqrt{\frac{7{,}514}{13{,}955{,}000}\,20(100-20)} = 0.9 \text{ percent}$$

The 90-percent confidence interval as shown is from 19 to 21 percent. Therefore, a conclusion that the average percentage derived from all possible samples lies within a range computed in this way would be correct for roughly 90 percent of all samples.

Standard error of a difference. The standard error of a difference between two sample estimates, x and y, is approximately equal to

$$s_{(x-y)} = \sqrt{s_x^2 + s_y^2 - 2rs_xs_y} \qquad (3)$$

where s_x and s_y are the standard errors of the estimates x and y and r is the correlation coefficient between the characteristics estimated by x and y. The estimates can be numbers, averages, percents, ratios, etc. Underestimates or overestimates of the standard error of differences result if the estimated correlation coefficient is overestimated or underestimated, respectively. In this report, r is assumed to be 0.

Illustration. Suppose we need the difference in the percentage of White children who lived with two parents and Black children who lived with two parents in 1991. From the 51,944,000 White children and the 10,571,000 Black children, 79 percent and 42 percent, respectively, lived with two parents in 1991. Using formula 2 and the "b" parameter, the standard errors of these percentages are approximately 0.5 percent and 1.3 percent, respectively.

Now, we compute the standard error of the difference using the above two standard errors. The correlation between these estimates is assumed to be zero. Therefore, we compute the standard error of the difference using formula 3.

$$s_{(x-y)} = \sqrt{(0.5)^2 + (1.3)^2} = 1.4 \text{ percent}$$

Suppose that it is desired to test at the 10-percent significance level whether the two percentages differ significantly. To perform the test, compare the difference of 37 percent to the product of 1.645 x 1.4 percent = 2.3 percent. Since the difference is larger than 1.645 times the standard error of the difference, the data show that the estimates of 79 and 42 percent differ significantly at the 10-percent level.

Table 3. **SIPP Topical Module Generalized Variance Parameters for the 1991 Panel**

Characteristics	a	b
Children 0 to 17 years	−0.0001340	7,514

Table 4. **Standard Errors of Estimated Numbers of Persons**

(Numbers in thousands)

Size of estimate	Standard error
200	39
300	47
500	61
1,000	86
2,000	120
3,000	146
5,000	185
7,500	221
10,000	248
15,000	287
25,000	323
35,000	314
40,000	294
45,000	258
48,000	228
50,000	202
52,000	168
55,000	89
60,000	24

Table 5. **Standard Errors of Estimated Percentages of Persons**

Base of estimated percentage (thousands)	Estimated percentages					
	≤ 1 or ≥ 99	2 or 98	5 or 95	10 or 90	25 or 75	50
200	1.93	2.71	4.22	5.81	8.39	9.69
300	1.57	2.22	3.45	4.75	6.85	7.91
500	1.22	1.72	2.67	3.68	5.31	6.13
1,000	0.86	1.21	1.89	2.60	3.75	4.33
2,000	0.61	0.86	1.34	1.84	2.65	3.06
3,000	0.50	0.70	1.09	1.50	2.17	2.50
5,000	0.39	0.54	0.84	1.16	1.68	1.94
7,500	0.31	0.44	0.69	0.95	1.37	1.58
10,000	0.27	0.38	0.60	0.82	1.19	1.37
15,000	0.22	0.31	0.49	0.67	0.97	1.12
25,000	0.17	0.24	0.38	0.52	0.75	0.87
35,000	0.15	0.21	0.32	0.44	0.63	0.73
40,000	0.14	0.19	0.30	0.41	0.59	0.69
45,000	0.13	0.18	0.28	0.39	0.56	0.65
48,000	0.12	0.18	0.27	0.38	0.54	0.63
50,000	0.12	0.17	0.27	0.37	0.53	0.61
52,000	0.12	0.17	0.26	0.36	0.52	0.60
55,000	0.12	0.16	0.25	0.35	0.51	0.58
60,000	0.11	0.16	0.24	0.34	0.48	0.56
65,000	0.11	0.15	0.23	0.32	0.47	0.54
66,000	0.11	0.15	0.23	0.32	0.46	0.53