APPENDIX

DEFINITIONS AND EXPLANATIONS

Population coverage. The data from the June 1975 survey are for the civilian population of the United States excluding the relatively small number of inmates of institutions.

Year of birth. The tables contain data on persons born in calendar years 1900 to 1959. Date of birth was obtained in terms of month and year.

For this and all other subjects covered in this report except family income, values were allocated to persons with nonresponses, through use in the electronic computer of the standard census procedure of substituting a value that was reported by a previously processed person of similar characteristics.

Age. The age classification is based on the age of the person at the last birthday. Age at first marriage was determined by subtracting the date of birth from the date of first marriage (all in months and years). Ages at other marital events were obtained by analogous procedures.

Race. The population is divided into three groups on the basis of race: White, Black, and other races. The last category includes Indians, Japanese, Chinese, and all other races except White and Black.

Spanish origin. Persons of Spanish origin in this report are those persons who indicated that their origin was Mexican, Puerto Rican, Cuban, Central or South American, or some other Spanish origin. The latter category includes persons from Spain as well as persons with combinations of types of Spanish origins. Persons who reported that they were of one of the specific Spanish origin categories and a non-Spanish category were included in the specific Spanish category. Persons of Spanish origin may be of any race.

Marital status. The marital status classification refers to the status at the time of enumeration. "Separated" is regarded as a subdivision of the category "married" and includes persons with legal separations, those living apart with intentions of obtaining a divorce, and other persons permanently or temporarily separated because of marital discord.

Marital history. In addition to their current marital status, persons 14 years old and over were asked how many times they had been married; when they had married for the first time; whether that marriage had ended in widowhood or divorce and when that marriage had ended (if it was no longer intact); and if they had remarried, they were asked when they had entered their latest marriage and when that marriage had ended (if it was no longer intact). All dates were recorded in terms of month and year, and this detail was used in deriving age at each event or interval between events.

Median. The median is the value which divides a distribution into two equal parts; one-half of the cases falling below this value and one-half of the cases exceeding this value.

SOURCE AND RELIABILITY OF THE ESTIMATES

Source of data. The estimates contained in these tables are based on data obtained from a supplement to the Current Population Survey (CPS) in June 1975. The CPS is spread over 461 areas comprising 923 counties and independent cities. These areas are chosen to provide coverage in each State and the District of Columbia. Approximately 47,000 occupied households are eligible for interview each month. Of this number, 2,000 occupied units, on the average, are visited but interviews are not obtained because the occupants are not found at home after repeated calls or are unavailable for some other reason. In addition to the 47,000, there are also about 8,000 sample units in an average month which are visited but are found to be vacant or otherwise not to be interviewed.

The CPS deals mainly with labor force data. Questions relating to labor force participation are asked about each member 14 years old or older in the household. In June 1975, additional questions relating to marital history were also asked of the same sample.

The estimation procedure used for the CPS data involves the inflation of the weighted sample results to independent estimates of the civilian noninstitutional population of the United States by age, race, and sex. These independent estimates were based on statistics from the 1970 Census of Population; statistics on births, deaths, immigration, and emigration; and statistics on the strength of the Armed Forces.

Reliability of the estimates. Since the estimates in these tables were based on a sample, they may differ somewhat from the figures that would have been obtained if a complete census had been taken using the same schedules, instructions, and enumerators. There are two types of errors possible in an estimate based on a sample survey—sampling and nonsampling. For estimates in this report, indications of the magnitude of sampling error are provided, but the extent of the nonsampling error is unknown. Consequently, particular care should be exercised in the interpretation of figures based on a relatively small number of cases or on small differences between estimates.

Nonsampling variability. As in any survey work, the results are subject to errors of response and nonreporting in addition to sampling variability. Nonsampling errors can be attributed to many sources, e.g., inability to obtain information about all cases in the sample, definitional difficulties, differences in the interpretation of questions, inability or unwillingness to provide correct information on the part of respondents, inability to recall information, mistakes made in collection such as in recording or coding the data, mistakes made in processing the data and mistakes made in estimating values for missing data. To date, emphasis has been placed on identification and control of nonsampling errors and not on providing estimates of magnitude of such errors in the data.

Sampling variability. The standard errors given in the tables are primarily measures of sampling variability, that is, of the variations that occur by chance because a sample rather than the whole of the population was surveyed. As calculated, the standard error also partially measures the effect of certain response and enumeration errors, but it does not measure any systematic biases in the data. The chances are about 68 out of 100 that an estimate from the survey differs from a complete census figure by less than the standard error. The chances are about 90 out of 100 that this difference would be less than 1.6 times the standard error, and chances are 95 out of 100 that the difference would be less than twice the standard error.

All the statements of comparison appearing in the text are significant at a 1.6 standard error level or better, and most are significant at a level of more than 2.0 standard errors. This means that for most differences cited in the text, the estimated difference is greater than twice the standard error of the difference. Statements of comparison qualified in some way (e.g., by use of the phrase, "some evidence") have a level of significance between 1.6 and 2.0 standard errors.

Note when using small estimates. Percent distributions are shown in the report only when the base of the percentage is 75,000 or greater. Because of the large standard errors involved, there is little chance that percentages would reveal useful information when computed on a smaller base. Estimated totals are shown, however, even though the relative standard errors of these totals are larger than those for corresponding percentages. These smaller estimates are provided primarily to permit such combinations of the categories as serve each user's needs.

Note on comparisons with data from other surveys. Data obtained from the Current Population Surveys and other surveys and sources are not entirely comparable, due in large part to differences in interviewer training and experience and in the differing survey processes. This is an additional component of error not reflected in the standard error tables; therefore, caution should be used in comparing results between these different sources.

Reliability of an estimated percentage. The reliability of an estimated percentage, computed by using sample data for both numerator and denominator, depends upon both the size of the percentage and the size of the total upon which the percentage is based. Estimated percentages are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more.

Standard error tables and their use. Instead of providing individual standard error tables for each characteristic of interest, generalized standard error tables for estimated numbers and estimated percentages, by race, are provided in tables A-1 through A-4 to conserve space.

The figures presented in these tables provide approximations to standard errors of various CPS estimates shown in this report. In all the standard error tables, standard errors for intermediate values not shown may be approximated by interpolation. In order to derive standard errors that would be applicable to a wide variety of items and could be prepared at a moderate cost, a number of approximations were required. In addition, where two or more items have nearly equal standard errors, such as total population and White population, one table is used to represent them. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than the precise standard error for any specific item.

Illustration of the use of tables of standard errors. Table B of this report shows that in June 1975 there were 6,597,000 widowed women in the population who had been married at least once. Table A-1 shows the standard error of an estimate of this size to be approximately 149,000. The chances are 68 out of 100 that the estimate would have been a figure differing from a complete census figure by less than 149,000. The chances are 95 out of 100 that the estimate would have been a figure differing from a complete census figure by less than 298,000 (twice the standard error).

Table A-1. Standard Errors of Estimated Numbers of Persons

Total or White Population[1]

(68 chances out of 100. Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
25	9	7,500	159
50	13	10,000	182
100	18	20,000	251
250	29	30,000	300
500	41	40,000	336
1,000	59	50,000	364
2,000	83	75,000	409
3,000	102	100,000	424
4,000	117	115,000	421
5,000	131		

[1]To obtain approximate standard errors of estimates of Spanish persons, multiply these standard errors by 1.46.

Table B also shows that of the 6,597,000 widowed women mentioned above, 15.9 percent had been married twice. Table A-3 shows that the standard error of 15.9 percent on a base of 6,597,000 is approximately 0.8. Consequently, the chances are 68 out of 100 that the estimated 15.9 percent would be within 0.8 percentage points of a complete census figure. Chances are 95 out of 100 that the estimate would be within 1.6 percentage points of a complete census figure, i.e., the 95 percent confidence interval would be from 14.3 to 17.5.

Table A-2. Standard Errors of Estimated Numbers of Persons

Black and Other Races

(68 chances out of 100. Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
25	11	3,000	115
50	15	4,000	129
100	22	5,000	141
250	35	7,500	161
500	49	10,000	171
1,000	69	15,000	167
2,000	95		

Standard error of a median. The standard error of an estimated median depends upon the form as well as on the size of the distribution from which the median is determined. An approximate method for measuring the reliability of a median is to determine an interval about the estimated median, such that there is a stated degree of confidence that the median based on a complete census lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. Using table A-3 or A-4, determine the standard error on a 50 percent characteristic, using the appropriate base[1].
2. Add to and subtract from 50 percent the standard error determined in step (1).
3. Using the distribution of the characteristic[1], calculate the confidence interval corresponding to the two points established in step (2).

A two standard error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step (1).

Standard error of a difference. For a difference between two sample estimates, the standard error is approximately equal to the square root of the sum of the squared standard errors of the estimates; the estimates can be of numbers, percents, ratios, medians, etc. This will represent the actual standard error quite accurately for the difference between two estimates of the same characteristic in two different areas, or for the difference between separate and uncorrelated characteristics in the same area. If, however, there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error.

Illustration of the computation of the standard error of a difference between estimated percentages. Table B of this report shows that of the 1,096,000 widowed males in June 1975, 16.1 percent had been married twice. The apparent difference between the 15.9 percent figure for females with the same marital history and the percentage of males is 0.2 percent. The standard error, σ_x, of the 15.9 percent is 0.8, as shown above. Table A-3 shows the standard error, σ_y, of 16.1 percent on a base of 1,096,000 to be approximately 2.1 percent. To get the standard error of the estimated difference, use the following formula:

$$\sigma_{(x-y)} = \sqrt{\sigma_x^2 + \sigma_y^2} \quad .$$

Therefore, the standard error of the difference of 0.2 percent is about

$$2.2 = \sqrt{(0.8)^2 + (2.1)^2} \quad .$$

This means the chances are 68 out of 100 that the estimated difference based on the sample estimates would vary from the difference derived using complete census figures by less than 2.2 percent. The 68 percent confidence interval around the 0.2 percent difference is from −2.0 to 2.4, i.e., 0.2 ± 2.2. A conclusion that the average estimate of the difference derived from all possible samples of the same size and design lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. The 95 percent confidence interval is −4.2 to 4.6. Thus, we cannot conclude with 95 percent confidence that there is a significant difference between the percentages for male and female, twice-married, widowed people.

[1] It was not feasible to print tables of distributions for the medians included in this report because of the amount of space required to do so. They will, however, be furnished upon request.

Table A-3. Standard Errors of Estimated Percentages of Persons

Total or White Population[1]

(68 chances out of 100)

Base of percentage (thousands)	Estimated percentage					
	1 or 99	2 or 98	5 or 95	10 or 90	25 or 75	50
100	1.9	2.6	4.1	5.6	8.1	9.4
250	1.2	1.7	2.6	3.5	5.1	5.9
500	0.8	1.2	1.8	2.5	3.6	4.2
1,000	0.6	0.8	1.3	1.8	2.6	3.0
2,500	0.4	0.5	0.8	1.1	1.6	1.9
5,000	0.3	0.4	0.6	0.8	1.1	1.3
10,000	0.2	0.3	0.4	0.6	0.8	0.9
25,000	0.14	0.2	0.3	0.4	0.5	0.6
50,000	0.08	0.14	0.2	0.3	0.4	0.4
100,000	0.06	0.08	0.15	0.2	0.3	0.3
120,000	0.05	0.08	0.14	0.2	0.2	0.3

[1] To obtain approximate standard errors of estimated percentages of Spanish persons, multiply these standard errors by 1.46.

Table A-4. Standard Errors of Estimated Percentages of Persons

Black and Other Races

(68 chances out of 100)

Base of percentage (thousands)	Estimated percentage					
	1 or 99	2 or 98	5 or 95	10 or 90	25 or 75	50
100	2.2	3.1	4.9	6.7	9.7	11.2
250	1.4	2.0	3.1	4.3	6.1	7.1
500	1.0	1.4	2.2	3.0	4.3	5.0
1,000	0.7	1.0	1.5	2.1	3.1	3.5
2,500	0.4	0.6	1.0	1.3	1.9	2.2
5,000	0.3	0.4	0.7	1.0	1.4	1.6
10,000	0.2	0.3	0.5	0.7	1.0	1.1
15,000	0.2	0.3	0.4	0.5	0.8	0.9