

Appendix

DEFINITIONS AND EXPLANATIONS

Population coverage. The figures shown are for the civilian population excluding [the relatively small number of] inmates of institutions.

Age. The age classification is based on the age of the person at his or her last birthday.

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 any other race except White and Black.

Preprimary school enrollment. Preprimary enrollment, as used here, is the sum of enrollment in regular kindergarten and nursery school.

A nursery school is defined as a group or class that is organized to provide educational experiences for children during the year or years preceding kindergarten. It includes instruction as an important and integral phase of its program of child care. Private homes in which essentially custodial care is provided are not considered nursery schools.

Full- and part-day attendance. Children attending nursery school were classified as attending during either part of the day or the full day. Part-day attendance refers to those who attend either in the morning or in the afternoon, but not both. Full-day attendance refers to those who attend both in the morning and afternoon.

Head Start. Children enrolled in Head Start programs or similar programs sponsored by local agencies to provide pre-school education to young children are counted under "nursery" or "kindergarten" as appropriate.

Mother. Many of the tables in this report present data on 3- to 5-year-old children living in households by characteristics of their "mother." Identification of the actual parents of children is not possible from October CPS data; however, relationship to the head of the household (i.e., related/not related) is ascertainable. Thus, the characteristics identified as those of the "mother" actually represent characteristics of the woman who was listed as the household head or as the wife of the household head. For the vast majority of children, this person is in fact their mother; however for some, this person may actually be a grandmother, aunt, or a non-relative (if two or more unrelated families live in the same household, for example).

Children in household. The greatest proportion (94.3 percent) of children under 14 years of age in households are actually children of the head of the household and thus may

be considered siblings. Other children in households include children who are related to the head in other ways, such as nieces, nephews, and grandchildren (4.7 percent) and children not related to the head (1.0 percent). In this report all children are treated as children of the head and siblings.¹

Labor force. Persons 16 years old and over are classified as in the labor force if they were employed as civilians, unemployed, or in the Armed Forces during the survey week. The "civilian labor force" is comprised of all civilians classified as employed or unemployed.

Employed. Employed persons comprise (1) all civilians who, during the specified week, did any work at all as paid employees or in their own business or profession, or on their own farm, or who worked 15 hours or more as unpaid workers on a farm or in a business operated by a member of the family; and (2) all those who were not working but who had jobs or businesses from which they were temporarily absent because of illness, bad weather, vacation, or labor-management dispute, or because they were taking time off for personal reasons, whether or not they were paid by their employers for time off, and whether or not they were seeking other jobs. Excluded from the employed group are persons whose only activity consisted of work around the house (such as own home housework, painting or repairing own home) or volunteer work for religious, charitable, and similar organizations.

Unemployed. Unemployed persons are those civilians who, during the survey week, had no employment but were available for work and (1) had engaged in any specific job-seeking activity within the past 4 weeks, such as registering at a public or private employment office, meeting with prospective employers, checking with friends or relatives, placing or answering advertisements, writing letters of application, or being on a union or professional register; (2) were waiting to be called back to a job from which they had been laid off; or (3) were waiting to report to a new wage or salary job within 30 days.

Not in the labor force. All civilians who are not classified as employed or unemployed are defined as "not in the labor force." This group who are neither employed nor seeking work includes persons engaged only in own home housework, attending school, or unable to work because of long-term physical or mental illness; persons who are retired or too old to work, seasonal workers for whom the survey week

¹ The proportions were obtained from Table 3 of *Current Population Reports*, Series P-20, No. 306, "Marital Status and Living Arrangements: March 1976," issued January 1977.

fell in an off season, and the voluntarily idle. Persons doing only unpaid family work (less than 15 hours during the reference week) are also classified as not in the labor force.

Full-time and part-time employment. Persons who worked 35 hours or more during the survey week and those who worked 1 to 34 hours but usually work full time are classified as employed full time. Part-time workers are persons who worked 1 to 34 hours during the survey week and usually work only 1 to 34 hours. Persons with a job but not at work during the survey week are classified according to whether they usually work full or part time.

Occupation of mother. The data refer to the civilian job held during the survey week. Persons employed at two or more jobs were reported in the job at which they worked the greatest number of hours during the week. The term "white collar" refers to the combination of the following major groups used in the 1970 Census of Population: (a) professional, technical and kindred; (b) managers and administrators, except farm; (c) clerical and kindred workers; and (d) sales workers. The term "blue collar" refers to the sum of the following major occupation groups: (a) craft and kindred workers; (b) operatives, including transport; and (c) laborers, except farm. The term "farm workers" includes the major occupation groups of (a) farmers and farm managers and (b) farm laborers and supervisors. The term "service workers" includes service workers, including private household.

Years of school completed. Data on years of school completed in this report were derived from the combination of answers to two questions: (a) "What is the highest grade of school he has ever attended?" and (b) "Did he finish this grade?"

The questions on educational attainment apply only to progress in "regular" schools. Such schools include graded public, private, and parochial elementary and high schools (both junior and senior high), colleges, universities, and professional schools, whether day schools or night schools. Thus, regular schooling is that which may advance a person toward

an elementary school certificate or high school diploma, or a college, university, or professional school degree. Schooling in other than regular schools was counted only if the credits obtained were regarded as transferable to a school in the regular school system.

Symbols. A dash "--" represents zero or a number which rounds to zero. The symbol "<" means less than. "B" means that the base is too small to show the derived measure; "NA" means not available; and "..." means not applicable. "S" means that the base for a derived number does not meet publication standards because the base is not comparable to CPS estimates.

Rounding of estimates. Individual figures are rounded to the nearest thousand without being adjusted to group totals, which are independently rounded. Percentages are based on the unrounded absolute numbers.

SOURCE AND RELIABILITY OF THE ESTIMATES

Source of data. The estimates in this report are based on data obtained annually in October of 1967 through 1976 from the Current Population Survey (CPS) conducted by the Bureau of the Census and from supplementary questions to the CPS. The monthly CPS deals mainly with labor force data. Questions relating to labor force participation are asked about each member 14 years or older in the household. In addition, supplementary questions are asked in October concerning educational characteristics, such as school enrollment, in order to acquire information about all levels of education.

The following table provides a description of some aspects of the CPS sample design in use during the referenced data collection periods.

Description of the Current Population Survey

Time period	Number of sample areas ¹	Households eligible		Households visited, not eligible ²
		Interviewed	Not interviewed	
Aug. 1972 to present.....	461	45,000	2,000	8,000
Aug. 1971 to July 1972.....	449	45,000	2,000	8,000
Jan. 1967 to July 1971.....	449	48,000	2,000	8,500

¹These areas were chosen to provide coverage in each State and the District of Columbia.

²These are households which were visited, but were found to be vacant or otherwise not eligible for interview.

The present CPS sample was initially selected from the 1970 census files and is updated continuously to reflect new construction where possible, (see section "Nonsampling Variability," below). Samples for previous sample designs were selected from files of the census most recently completed at the time.

The estimation procedure used for both the CPS and supplemented 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 this report are based on a sample, they may differ somewhat from the figures that would have been obtained had a complete census been taken using the same schedule, instructions and enumerators. There are two types of errors possible in an estimate based on a sample survey—sampling and nonsampling. The standard errors provided for this report primarily 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 full 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, mistakes made in estimating values for missing data, and failure to represent all units within the sample (undercoverage).

The approximate magnitude of two sources of undercoverage in CPS is known. About 600,000 conventional new construction units (housing units other than mobile homes) were assigned building permits prior to the 1970 census but building was not completed by the time of the census (i.e., April 1970); these units have no representation in the CPS sample. Most conventional new construction, for which building permits were issued after the census, is represented. About 290,000 occupied mobile homes are not represented in CPS; these units were either missed in the census or have been built or occupied since the census. The extent of other sources of undercoverage is unknown. Note that these estimates of missed units are relevant to the present sample

only and not to earlier designs where the extent of undercoverage was generally less. The estimation procedure described previously does partially adjust for undercoverage through the ratio estimation to independent population controls. However, if for particular characteristics, the persons occupying the missed households differ from those occupying sampled households, then for these characteristics some nonsampling error would still result from the undercoverage.

Sampling variability. The standard errors given in the following tables are primarily measures of sampling variability, that is, of the variations that occurred by chance because a sample rather than the whole of the population was surveyed. 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 about 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 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 numbers are shown, however, even though the relative standard errors of these numbers 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.

Standard error tables and their use. In order to derive standard errors that would be applicable to a large number of estimates and could be prepared at a moderate cost, a number of approximations were required. Therefore, instead of providing an individual standard error for each estimate, generalized sets of standard errors are provided for various types of characteristics. As a result, the sets of standard errors provided give an indication of the order of magnitude of the standard error of an estimate rather than the precise standard error.

The figures presented in tables A-1 and A-2 are approximations to standard errors of estimates for all children in nursery schools or kindergartens in the total U.S. Estimated standard errors for specific characteristics cannot be obtained from tables A-1 and A-2 without the use of the factors in table A-3. These factors must be applied to the generalized standard errors in order to adjust for the effects of sampling

Table A-1. Standard Errors of Estimated Numbers

(68 chances out of 100. Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
1.....	1	750.....	35
10.....	4	1,000.....	39
25.....	7	1,500.....	47
50.....	9	2,500.....	57
100.....	13	5,000.....	65
250.....	21	7,500.....	54
500.....	29		

Table A-2. Standard Errors of Estimated Percentages

Estimated percentage	Base of estimated percentage (thousands)									
	75	100	250	500	750	1,000	1,500	2,500	5,000	7,500
2 or 98.....	2.1	1.9	1.2	0.8	0.7	0.6	0.5	0.4	0.3	0.2
5 or 95.....	3.3	2.9	1.8	1.3	1.0	0.9	0.7	0.6	0.4	0.3
10 or 90.....	4.6	4.0	2.5	1.8	1.4	1.3	1.0	0.8	0.6	0.5
20 or 80.....	6.1	5.3	3.3	2.4	1.9	1.7	1.4	1.1	0.7	0.6
35 to 65.....	7.3	6.3	4.0	2.8	2.3	2.0	1.6	1.3	0.9	0.7
50.....	7.6	6.6	4.2	2.9	2.4	2.1	1.7	1.3	0.9	0.7

Table A-3. "f" Factors to be Applied to Tables A-1 and A-2 to Approximate Standard Errors

Type of characteristic	Value of f
Kindergarten and nursery school enrollment.....	1.0
Labor force status of mother.....	1.1
Mother unemployed.....	1.1
Marital status of mother.....	1.0
Number of siblings.....	1.4
Years of school completed by mother.....	1.1

Table A-4. Parameters for Direct Computation of Standard Errors of Estimated Numbers and Percentages of Children

Type of characteristic	Parameters	
	a	b
KINDERGARTEN AND NURSERY SCHOOL ENROLLMENT		
All Children		
3 to 5 years old.....	-0.000179	1738
3 and 4 years old.....	-0.000279	1738
3 years old.....	-0.000576	1738
4 years old.....	-0.000540	1738
5 years old.....	-0.000498	1738
White Children		
3 to 5 years old.....	-0.000217	1738
3 and 4 years old.....	-0.000339	1738
5 years old.....	-0.000603	1738
Black Children		
3 to 5 years old.....	-0.001160	1738
3 and 4 years old.....	-0.001800	1738
5 years old.....	-0.003260	1738
LABOR FORCE STATUS OF MOTHER		
(Employment, occupation, and not in labor force)		
All Children		
3 to 5 years old.....	-0.000214	2078
3 and 4 years old.....	-0.000333	2078
3 years old.....	-0.000689	2078
4 years old.....	-0.000645	2078
5 years old.....	-0.000596	2078
White Children		
3 to 5 years old.....	-0.000259	2078
3 and 4 years old.....	-0.000405	2078
Black Children		
3 to 5 years old.....	-0.001390	2078
3 and 4 years old.....	-0.002150	2078
Mother unemployed.....	-0.000203	1971
MARITAL STATUS OF MOTHER		
All Children		
3 to 5 years old.....	-0.000179	1738
3 and 4 years old.....	-0.000279	1738
5 years old.....	-0.000498	1738
NUMBER OF SIBLINGS		
All Children		
3 to 5 years old.....	-0.000360	3500
3 and 4 years old.....	-0.000561	3500
5 years old.....	-0.001000	3500
YEARS OF SCHOOL COMPLETED BY MOTHER		
All Children		
3 to 5 years old.....	-0.000212	2064
3 and 4 years old.....	-0.000331	2064
3 years old.....	-0.000684	2064
4 years old.....	-0.000641	2064
5 years old.....	-0.000592	2064

design and estimating procedure on the value of the characteristic. Standard errors for intermediate values not shown in the generalized tables may be approximated by interpolation.

Two parameters (denoted "a" and "b") are used to calculate standard errors for each type of characteristic; they are presented in table A-4. These parameters were used to calculate the entries in tables A-1 and A-2, and to calculate the factors in table A-3. They also may be used to calculate directly the standard errors for estimated numbers and estimated percentages. Direct computation of the standard errors will give more accurate results than the use of the standard error tables and the factors in table A-3. Methods for direct computation are given in the following sections.

Standard errors of estimated numbers. The approximate standard error, σ_x , of an estimated number shown in this report can be obtained in two ways. It may be obtained by use of the formula

$$\sigma_x = f\sigma \quad (1)$$

where f is the appropriate factor from table A-3, and σ is the standard error of the estimate obtained by interpolation from table A-1. Alternatively, standard errors may be approximated by the following formula, (2), from which the standard errors were calculated in table A-1. Use of this formula will provide more accurate results than the use of formula (1) above.

$$\sigma_x = \sqrt{ax^2 + bx} \quad (2)$$

Here x is the size of the estimate and a and b are the parameters in table A-4 associated with the particular type of characteristic.

Standard errors of estimated percentages. The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends on both size of the percentage and the size of the total upon which this 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. When the numerator and denominator of the percentage are in different categories, use the factor or parameters indicated by the numerator. The approximate standard error, $\sigma_{(x,p)}$, of an estimated percentage can be obtained by use of the formula

$$\sigma_{(x,p)} = f\sigma \quad (3)$$

In this formula f is the appropriate factor from table A-3, and σ is the standard error on the estimate from table A-2. Alternatively, standard errors may be approximated by the following formula, (4), from which the standard errors in table A-2 were calculated; direct computation will give more accurate results than use of the standard error tables and the factors.

$$\sigma_{(x,p)} = \sqrt{\frac{b}{x} \cdot p(100-p)} \quad (4)$$

Here x is the size of the subclass of persons which is the base of the percentage, p is the percentage ($0 \leq p \leq 100$), and b is the parameter in table A-4 associated with the particular type of characteristic in the numerator of the percentage.

Illustration of the use of standard error tables. Table 1 of this report shows that in 1976 there were 4,059,000 3- to 5-year-old children whose mothers were in the labor force. Table A-1 shows the standard error of an estimate of this size to be approximately 62,000. Applying the appropriate factor from table A-3 and using formula (1), the approximate standard error is $1.1 \times 62,000 = 68,000$. The chances are 68 out of 100 that the estimate would have been a figure differing from a complete census figure by less than 68,000. The chances are 95 out of 100 that the estimate would have differed from a complete census figure by 136,000 (twice the standard error). Alternatively, using formula (2) and the parameters $a = .000214$ and $b = 2078$ from table A-4 gives an estimate of the standard error of 70,000.

Table 1 also shows that of these 4,059,000 3-5 year olds with mothers in the labor force 2,136,000 or 52.6 percent of them are enrolled in school. The numerator of this percentage is children enrolled in school whose mothers are in the labor force. This characteristic relates to two categories in table A-4: enrollment in school and labor force status of the mother. Since the larger b -parameter should be used whenever the characteristic relates to more than one b -parameter, the proper b is 2078. Using formula (4), the standard error of an estimate of 52.6 percent is:

$$\sigma_{(x,p)} = \sqrt{\frac{2078}{4059000} (52.6) (100-52.6)} = 1.1 \text{ percent}$$

Consequently, the chances are 68 out of 100 that the estimated 52.6 percent would be within 1.1 percentage points of a complete census figure. Chances are 95 out of 100 that the estimate would be within 2.2 percentage points of a complete census figure. As an alternative, tables A-2 and A-3 can be used to get an estimated standard error of $1.1 \times 1.1 = 1.2$ percent on the estimate of 52.6 percent.

Standard error of a difference. For a difference between two sample estimates, the standard error is approximately equal to

$$\sigma_{(x-y)} \doteq \sqrt{\sigma_x^2 + \sigma_y^2} \quad (5)$$

Where σ_x and σ_y are the standard error of the estimates x and y ; the estimate can be numbers, percents, averages, 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 two separate and uncorrelated characteristics in the same areas. 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. Table 1 also shows that of the 5,496,000 children with mothers not in the labor force, 47.1 percent are enrolled in schools. Using formula (4), the standard error for

this estimate is 1.0 percent. It was shown above that 52.6 percent of children with mothers in the labor force are enrolled in school and the standard error on this estimate is 1.1 percent. Therefore, using formula (5), the standard error on the estimated difference of 5.5 percent is about

$$1.5 \doteq \sqrt{(1.0)^2 + (1.1)^2}$$

This means that the chances are 68 out of 100 that the estimated difference based on the sample estimate would vary from the difference derived using complete census figures by less than 1.5 percent. The 68 percent confidence interval about the 5.5 percent difference is from 4.0 percent to 7.0 percent, i.e., 5.5 ± 1.5 . A conclusion that the average estimate of the difference derived over 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 from 2.5 to 8.5. Thus we can conclude with 95 percent confidence that there is a difference in the school enrollment rates of 3- to 5-year-old children with mothers in the labor force and the rates for children with mothers not in the labor force.