Appendix. Source and Reliability of Estimates

SOURCE OF DATA

The estimates of school enrollment in 1984 are based on data obtained in October 1984 from the Current Population Survey (CPS) conducted by the Bureau of the Census. The CPS sample was initially selected from the 1970 census file and is updated continuously to reflect new construction where possible. The sample is composed of approximately 61,500 occupied households that are eligible for interview. Of this number, about 2,500 occupied units were visited, but interviews were not obtained because the occupants were not found at home after repeated calls or were unavailable for some other reason.

The estimation procedure used for this survey involved 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 are based on statistics from decennial censuses; statistics on births, deaths, immigration, and emigration; and statistics on the strength of the Armed Forces.

Data from 1972 to 1980 were obtained using independent population estimates based on the 1970 decennial census. Estimates for earlier years were based on earlier censuses. Data for 1981 and later years utilized independent estimates derived from the 1980 decennial census.

Two sets of estimates for 1981 are shown in some of the tables in this report; one set results from using independent population estimates based on the more up-to-date 1980 decennial census, and the other set results from using 1970 census-based population estimates. The 1970-based estimates have been included to provide continuity in the time series with previous years. Comparing the 1980-based estimates with the 1970-based estimates provides a measure of the effect of changing to the 1980-based estimation procedure.

CPS Design Phase-In. Since the inception of the CPS in 1940, the sample has been redesigned several times, most recently in the early 1970's, to upgrade the quality and reliability of the data and to meet changing data needs. Beginning in April 1984, the CPS design was being phased out through a series of changes that were completed in July 1985. The October 1984 CPS sample consisted of four rotation groups where sample segments were obtained by sampling from 1970 census materials, and four rotation groups where sample segments in continuing areas were obtained by sampling from 1980 census materials and segments in outgoing areas were based on 1970 census materials. A continuing area was one that stayed in sample with both the 1970 and 1980 designs, and an outgoing area was one that was no longer in sample in the 1980 design. The October sample had five rotation groups located in 629 sampling areas comprising 1,148 counties, independent cities, and minor civil divisions and three rotation groups located in 729 sample areas representing 1,973 counties and equivalent geographic areas. The coverage was in all 50 States and the District of Columbia.

For a description of the previous CPS sample designs, see the detailed reports in the series, Current Population Reports, Series P-60, Bureau of the Census, U.S. Department of Commerce.

RELIABILITY OF ESTIMATES

Since the estimates in this report are 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 questionnaires, 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 they 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.

Sampling Variability. The standard errors presented in tables A-1 and A-2 are primarily measures of sampling variability; that is, of the variations that occurred by chance because a sample rather than the entire population was surveyed. The sample estimates and its standard error enable one to construct interval estimates that include the average results of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under identical conditions using the same design and an estimate and its standard error were calculated from each sample, then:

1. Approximately 90 percent of the intervals from 1.6 standard errors below the estimates to 1.6 standard errors above the estimate would include the average results of all possible samples.

Table A-1. Standard Errors of Estimated Numbers of Persons 3 to 34 Years Old Enrolled in School, for the Total, Black, and Spanish-Origin Population: October 1984

(Numbers in thousands)

Enrollment	Total		Black		Spanish origin	
	Persons	Standard error	Persons	Standard error	Persons	Standard error
Total enrolled.....	57,313	265	8,226	113	4,284	203
Nursery.....	2,354	78	340	30	117	18
Private.....	1,593	65	161	21	39	10
Kindergarten.....	3,484	94	563	37	293	28
Private.....	531	38	51	12	26	8
Elementary.....	26,838	204	4,123	67	2,384	78
Private.....	2,718	84	177	22	166	21
High school.....	13,777	163	2,061	67	966	61
College.....	10,859	148	1,138	52	524	40
Full time.....	7,822	128	810	44	356	31

Note: Controlled to 1980 census base.

Source: Estimates from table 1.

Table A-2. Standard Errors of Estimated Percentages of Persons 3 to 34 Years Old Enrolled in School, for the Total, Black, and Spanish-Origin Population: October 1984

(Numbers in thousands)

Age	Total		Black		Spanish origin	
	Percentage	Standard error	Percentage	Standard error	Percentage	Standard error
3 to 34 years.....	47.9	0.2	50.1	0.6	47.7	1.0
3 and 4 years.....	36.3	0.9	38.2	2.4	24.2	3.4
5 and 6 years.....	94.5	0.5	94.1	1.2	93.9	1.9
7 to 9 years.....	99.0	0.2	99.5	0.3	98.7	0.8
10 to 13 years.....	99.4	0.11	99.3	0.3	99.4	0.5
14 and 15 years.....	97.8	0.3	97.9	0.7	94.9	1.9
16 and 17 years.....	91.5	0.5	92.4	1.3	85.7	3.0
18 and 19 years.....	50.1	0.9	44.3	2.4	39.9	4.1
20 and 21 years.....	33.9	0.8	27.7	2.2	28.1	3.9
22 to 24 years.....	17.3	0.5	15.7	1.4	11.3	2.0
25 to 29 years.....	9.1	0.3	7.4	0.8	7.1	1.3
30 to 34 years.....	6.3	0.3	6.5	0.8	5.2	1.3

Note: Controlled to 1980 census data.

Source: Estimates from table 6.

2. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average results of all possible samples.

The average result of all possible samples may or may not be contained in any particular computed interval. However, for a particular sample one can say with a specified confidence that the average results of all possible samples is included within the constructed interval.

All statements of comparison in the text have passed a hypothesis test at the 0.10 level of significance or better, and most have passed a hypothesis test at the 0.05 level of significance or better. This means that, for most differences cited in the text, the difference between two sample estimates 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 this report only when the base of the percentage is greater than 75,000. 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 of persons are shown, however, even though the relative standard errors of these numbers are larger than those for the corresponding percentages. These smaller estimates are provided primarily to permit those combinations of the categories which serves each user's needs.

Standard Errors for Data Based on the CPS Sample. Since this is an advance report, standard errors are provided in tables A-1 and A-2, and standard error parameters are provided in table A-3 for estimated numbers of persons and estimated percentages for only certain characteristics which are considered the most important among the data in the report. A more complete source and reliability statement for the 1984 data will be published with the forthcoming 1984 detailed report.

Standard Errors of Estimated Numbers and Estimated Percentages. The approximate standard errors of estimated numbers and percentages can be computed directly with formulas (1) and (2) below, respectively. The formulas are:

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

where "x" is the size of the estimate and "a" and "b" are the parameters associated with the characteristic; and

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

Table A-3. Parameters to be Used for School Enrollment Characteristics for Direct Computation of Standard Errors (Use for 1980 and Later Years)

Characteristic	Parameters[1]	
	a	b
Persons enrolled in school		
3 to 34 years old:		
Total or White	-0.000019	2,312
Black	-0.000129	2,600
Spanish origin (level)	+0.001744	2,131
Spanish origin (percentage only)	(X)	[2]3,873
14 to 34 years old:		
Total or White	-0.000028	2,312
Black	-0.000195	2,600
Spanish origin (level)	+0.001174	2,131
Spanish origin (percentage only)	(X)	[2]3,873
3 to 13 years old:		
Total, White or Spanish origin	-0.000073	2,698
Black	-0.000393	2,698
Spanish origin (percentage only)	(X)	[2]3,873

X - Not applicable

[1]To obtain "a" and "b" parameters for the October 1967-1979 supplements, multiply these parameters by 0.871. To obtain "a" and "b" parameters for 1965, multiply these parameters by 1.307.

[2]This "b" parameter is to be used to calculate the standard error of percentages only.

where "x" is the size of the subclass of the population which is the base of the percentage, "p" is the percentage ($0 < p < 100$), and "b" is the parameter associated with the characteristic.

Table A-3 provides the values of the "a" and "b" parameters that are used in formulas (1) and (2) to approximate standard errors of estimated numbers of persons and estimated percentages.

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

$$\sigma_{(x-y)} \doteq \sqrt{\sigma_x^2 + \sigma_y^2} \qquad (3)$$

where σx and σy are the standard errors of the estimates x and y; the estimates can be numbers, percents, ratios, etc. This will represent the estimated 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 area. If however, there is a high positive (negative) correlation between the two characteristics, the formula will overestimate (underestimate) the true standard error.