Attachment C

Table of Contents

Tables

Source and Accuracy of the Data for the 2005 Annual Social and Economic Supplement Microdata File

SOURCES OF DATA

The data in this microdata file come from the 2005 Annual Social and Economic Supplement (ASEC). The Census Bureau conducts the ASEC over a three-month period, in February, March, and April, with most data collection occurring in the month of March. The ASEC uses two sets of questions: the basic Current Population Survey (CPS) and a set of supplemental questions. The CPS, sponsored jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics, is the country's primary source of labor force statistics for the entire population. The U.S. Census Bureau and the U.S. Bureau of Labor Statistics also jointly sponsor the ASEC.

Basic CPS. The monthly CPS collects primarily labor force data about the civilian noninstitutional population living in the United States. Interviewers ask questions concerning labor force participation about each member 15 years old and over in sample households.

The CPS uses a multistage probability sample based on the results of the decennial census. When files from the most recent decennial census become available, the Census Bureau gradually introduces a new sample design for the CPS[1].

In April 2004, the Census Bureau began phasing out the 1990 sample and replacing it with the 2000 sample, creating a mixed sampling frame. Two simultaneous changes occured during this phase-in period. First, primary sampling units (PSUs)[2] selected for only the 2000 design gradually replaced those selected for the 1990 design. This involved 10 percent of the sample. Second, within PSUs selected for both the 1990 and 2000 designs, sample households from the 2000 design gradually replaced sample households from the 1990 design. This involved about 90 percent of the entire sample. By July 2005, the new sample design was completely implemented, and the sample came entirely from Census 2000 files.

In the first stage of the sampling process, PSUs are selected for sample. In the 1990 design, the United States was divided into 2,007 PSUs. These were then grouped into 754 strata, and one PSU was selected for sample from each stratum. In the 2000 sample design, the United States is divided into 2,025 PSUs. These PSUs are then grouped into 824 strata. Within each stratum, a single PSU is chosen for the sample, with its probability of selection proportional to its population as of the most recent decennial census. This PSU represents the entire stratum from which it was selected. In the case of strata consisting of only one PSU, the PSU is chosen with certainty.

The 1990 design and 2000 design stratum numbers are not directly comparable, since the 1990 design contained some PSUs in New England and Hawaii that were based on minor civil divisions instead of counties while the PSUs in the 2000 design are strictly county-based. The

---

[1] For detailed information on the 1990 sample redesign, see the Department of Labor, Bureau of Labor Statistics report, *Employment and Earnings*, Volume 41 Number 5, May 1994.

[2] The PSUs correspond to substate areas, counties, or groups of counties that are geographically contiguous.

PSUs have also been redefined to correspond to the new Office of Management and Budget (OMB) definitions of Core-Based Statistical Area definitions and to improve efficiency in field operations.

Approximately 72,700 households were selected for sample from the mixed sampling frame in March. Based on eligibility criteria, 11 percent of these households were sent directly to Computer-Assisted Telephone Interviewing (CATI). The remaining units were assigned to interviewers for Computer-Assisted Personal Interviewing (CAPI).[3] Of all housing units in sample, about 60,100 were determined to be eligible for interview. Interviewers obtained interviews at about 54,400 of these units. Noninterviews occur when the occupants are not found at home after repeated calls or are unavailable for some other reason. Table 1 summarizes changes in the CPS designs for the years in which data appear in this report.

The Annual Social and Economic Supplement. In addition to the basic CPS questions, interviewers asked supplementary questions for the ASEC. They ask these questions of the civilian noninstitutional population and also of military personnel who live in households with at least one other civilian adult. The additional questions cover the following topics:

- Household and Family Characteristics
- Marital Status
- Geographic Mobility
- Foreign Born Population
- Income from the previous calendar year
- Poverty
- Work Status/Occupation
- Health Insurance Coverage
- Program Participation
- Educational Attainment

Including the basic CPS sample, approximately 98,700 housing units are in sample for the ASEC. About 84,700 are determined to be eligible for interview and about 77,200 interviews are obtained (see Table 1).

The additional sample for the ASEC provides more reliable data for Hispanic households, non-Hispanic minority households, and non-Hispanic White households with children 18 years or younger. These households were identified for sample from previous months and the following April. For more information about the households eligible for the ASEC, please refer to:

> Technical Paper 63RV, *Current Population Survey: Design and Methodology*, U.S. Census Bureau, U.S. Department of Commerce, 2002. (http://www.census.gov/prod/2002pubs/tp63rv.pdf)

Table 1. Description of the of the March CPS Sample Cases: Basic + ASEC

[3] For further information on CATI and CAPI and the eligibility criteria, please see: Technical Paper 63RV, *Current Population Survey: Design and Methodology*, U.S. Census Bureau, U.S. Department of Commerce, 2002. (http://www.census.gov/prod/2002pubs/tp63rv.pdf)

Time Period	Number of Sample Areas	Basic CPS Housing Units Eligible		Total (ASEC + Basic CPS [1]) Housing Units Eligible	
		Interviewed	Not Interviewed	Interviewed	Not Interviewed
2005	754/824 [2]	54,400	5,700	77,200	7,500
2004	754	55,000	5,200	77,700	7,000
2003	754	55,500	4,500	78,300	6,800
2002	754	55,500	4,500	78,300	6,600
2001	754	46,800	3,200	49,600	4,300
2000	754	46,800	3,200	51,000	3,700
1999	754	46,800	3,200	50,800	4,300
1998	754	46,800	3,200	50,400	5,200
1997	754	46,800	3,200	50,300	3,900
1996	754	46,800	3,200	49,700	4,100
1995	792	56,700	3,300	59,200	3,800
1990 to 1994	729	57,400	2,600	59,900	3,100
1989	729	53,600	2,500	56,100	3,000
1986 to 1988	729	57,000	2,500	59,500	3,000
1985	629/729 [3]	57,000	2,500	59,500	3,000
1982 to 1984	629	59,000	2,500	61,500	3,000
1980 to 1981	629	65,500	3,000	68,000	3,500
1977 to 1979	614	55,000	3,000	58,000	3,500
1976	624	46,500	2,500	49,000	3,000
1973 to 1975	461	46,500	2,500	49,000	3,000
1972	449/461 [4]	45,000	2,000	45,000	2,000
1967 to 1971	449	48,000	2,000	48,000	2,000
1963 to 1966	357	33,400	1,200	33,400	1,200
1960 to 1962	333	33,400	1,200	33,400	1,200
1959	330	33,400	1,200	33,400	1,200

Notes: 1) The ASEC was referred to the Annual Demographic Survey (ADS) until 2002.
2) The Census Bureau redesigned the CPS following the Census 2000. During phase-in of the new design, housing units from the new and old designs were in the sample.
3) The Census Bureau redesigned the CPS following the 1980 Decennial Census of Population and Housing.
4) The Census Bureau redesigned the CPS following the 1970 Decennial Census of Population and Housing.

Estimation Procedure. This survey's estimation procedure adjusts weighted sample results to agree with independently derived population estimates of the civilian noninstitutional population of the United States. The adjusted estimate is called the post-stratification ratio estimate. The population estimates, used as controls for the CPS, are prepared annually to agree with the most current set of population estimates that are released as part of the Census Bureau's population estimates and projections program.

The population controls for the nation are distributed by demographic characteristics in two ways:

- Age, sex, and race (White alone, Black alone, Asian alone, and all other groups combined), and
- Age, sex, and Hispanic origin.

The projections for the states are distributed by race (Black alone and all other race groups combined), age (0-15, 16-44, and 45 and over), and sex.

The independent estimates by age, sex, and race, and Hispanic origin and for states by selected age groups and broad race categories are developed using the basic demographic accounting formula whereby the population from the latest decennial data is updated using data on the components of population change (births, deaths, and net international migration) with internal migration as an additional component in the state population estimates.

The net international migration component in the population estimates includes a combination of:

- Legal migration to the United States,
- Emigration of foreign-born and native people from the United States,
- Net movement between the United States and Puerto Rico,
- Estimates of temporary migration, and
- Estimates of net residual foreign-born population, which include unauthorized migration.

Because the latest available information on these components lag the survey date, it is necessary to make short-term projections of these components to develop the estimate for the survey date.

The estimation procedure of the ASEC included a further adjustment so husband and wife of a household received the same weight.

ACCURACY OF ESTIMATES

A sample survey estimate has two types of error: sampling and nonsampling. The accuracy of an estimate depends on both types of error. The nature of the sampling error is known given the survey design; the full extent of the nonsampling error is unknown.

Sampling Error. Since the CPS estimates come from a sample, they may differ from figures from an enumeration of the entire population using the same questionnaires, instructions, and enumerators. For a given estimator, the difference between an estimate based on a sample and the estimate that would result if the sample were to include the entire population is known as sampling error. Standard errors, as calculated by methods described in “Standard Errors and their Use,” are primarily measures of the magnitude of sampling error. However, they may include some nonsampling error.

Nonsampling Error. For a given estimator, the difference between the estimate that would result if the sample were to include the entire population and the true population value being estimated is known as nonsampling error. Sources of nonsampling errors include the following:

- Inability to obtain information about all cases in the sample (nonresponse)
- Definitional difficulties
- Differences in the interpretation of questions
- Respondent inability or unwillingness to provide correct information
- Respondent inability to recall information
- Errors made in data collection, such as in recording or coding the data
- Errors made in processing the data
- Errors made in estimating values for missing data
- Failure to represent all units with the sample (undercoverage).

Answers to questions about money income often depend on the memory or knowledge of one person in a household. Recall problems can cause underestimates of income in survey data, because it is easy to forget minor or irregular sources of income. Respondents may also misunderstand what the Census Bureau considers money income or may simply be unwilling to answer these questions correctly because the questions are considered too personal. See Appendix C, Current Population Reports, Series P60-184, *Money Income of Households, Families, and Persons in the United States: 1992* for more details.

To minimize these errors, the Census Bureau employs quality control procedures in sample selection, wording of questions, interviewing, coding, data processing, and data analysis.

Two types of nonsampling error that can be examined to a limited extent are nonresponse and undercoverage.

Nonresponse. The effect of nonresponse cannot be measured directly, but one indication of its potential effect is the nonresponse rate. For the cases eligible for the 2005 ASEC, the basic CPS nonresponse rate was 9.4 percent. The nonresponse rate for the Annual Social and Economic Supplement was an additional 8.8 percent. These two nonresponse rates lead to a combined supplement nonresponse rate of 17.4 percent.

Coverage. The concept of coverage in the survey sampling process is the extent to which the total population that could be selected for sample "covers" the survey's target population. CPS undercoverage results from missed housing units and missed people within sample households. Overall CPS undercoverage for March 2005 is estimated to be about 10 percent. CPS undercoverage varies with age, sex, and race. Generally, undercoverage is larger for males than for females and larger for Blacks than for Non-Blacks.

The CPS weighting procedure partially corrects for bias due to undercoverage, but biases may still be present when people who are missed by the survey differ from those interviewed in ways other than age, race, sex, Hispanic ancestry, and state of residence. How this weighting procedure affects other variables in the survey is not precisely known. All of these considerations affect comparisons across different surveys or data sources.

A common measure of survey coverage is the coverage ratio, calculated as the estimated population before post-stratification divided by the independent population control. Table 2 shows March 2005 CPS coverage ratios for certain age-sex-race-ancestry groups. The CPS coverage ratios can exhibit some variability from month to month.

Table 2. CPS Coverage Ratios : March 2005											
	Totals			White Only		Black Only		Residual Race		Hispanic	
Age Group	All People	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
0-15	0.92	0.92	0.92	0.94	0.94	0.81	0.78	0.95	0.98	0.97	0.94
16-19	0.88	0.90	0.85	0.91	0.88	0.78	0.71	0.97	0.94	1.03	0.94
20-24	0.81	0.80	0.82	0.82	0.84	0.59	0.72	0.91	0.76	0.83	0.84
25-34	0.84	0.81	0.87	0.84	0.89	0.66	0.79	0.82	0.86	0.76	0.87
35-44	0.89	0.86	0.93	0.88	0.95	0.70	0.80	0.85	0.88	0.84	0.94
45-54	0.91	0.89	0.93	0.90	0.94	0.80	0.85	0.88	0.96	0.81	0.91
55-64	0.91	0.91	0.90	0.91	0.91	0.86	0.89	0.90	0.83	0.88	0.82
65+	0.94	0.95	0.93	0.96	0.94	0.94	0.95	0.90	0.83	0.78	0.89
15+	0.89	0.88	0.90	0.89	0.92	0.75	0.82	0.88	0.87	0.83	0.90
0+	0.90	0.89	0.91	0.90	0.92	0.77	0.81	0.89	0.90	0.87	0.91

Notes: (1) The Residual Race group includes cases indicating a single race other than White or Black, and cases indicating two or more races.
(2) Hispanics may be of any race.

Comparability of Data. Data obtained from the CPS and other sources are not entirely comparable. This results from differences in interviewer training and experience and in differing survey processes. This is an example of nonsampling variability not reflected in the standard errors. Therefore, caution should be used when comparing results from different sources.

Caution should also be used when comparing data from this microdata file, which reflects Census 2000-based population controls, with microdata files from March 1994-2001, which reflect 1990 census-based population controls, and with microdata files from earlier years. Microdata files from previous years reflect the latest available census-based population controls. Be sure to compare data from microdata files with the same controls when possible. Although this change in population controls has relatively little impact on summary measures, such as averages, medians, and percentage distributions, it does have a significant impact on levels. For example, use of Census 2000-based population controls results in about a one percent increase in the civilian noninstitutional population and in the number of families and households. Thus, estimates of levels for data collected in 2002 and later years will differ from those for earlier years by more than what could be attributed to actual changes in the population. These differences could be disproportionately greater for certain population subgroups than for the total population.

Caution should also be used when comparing Hispanic estimates over time. No independent population control totals for people of Hispanic ancestry were used before 1985.

Users should also exercise caution due to changes caused by the phase-in of the Census 2000 files. During this time period, CPS data are collected from sample designs based on different censuses. Three features of the new CPS design have the potential of affecting published estimates: (1) the temporary disruption of the rotation pattern from August 2004 through June 2005 for a comparatively small portion of the sample, (2) the change in sample areas, and (3) the introduction of the new Core-Based Statistical Areas (formerly called metropolitan area). Most of the known effect on estimates during and after the sample redesign will be the result of changing from 1990 to 2000 geographic definitions. Research has shown that the national-level estimates of the metropolitan and nonmetropolitan populations should not change appreciably because of the new sample design. However, users should still exercise caution when comparing metropolitan and nonmetropolitan estimates across years with a design change, especially at the state level.

A Nonsampling Error Warning. Since the full extent of the nonsampling error is unknown, one should be particularly careful when interpreting results based on small differences between estimates. Even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test. Caution should also be used when interpreting results based on a relatively small number of cases. Summary measures (such as medians and percentage distributions) probably do not reveal useful information when computed on a subpopulation smaller than 75,000.

For additional information on nonsampling error including the possible impact on CPS data when known, refer to

- Statistical Policy Working Paper 3, *An Error Profile: Employment as Measured by the Current Population Survey*, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce, 1978. (http://www.fcsm.gov/working-papers/spp.html)

- Technical Paper 63RV, *Current Population Survey: Design and Methodology*, U.S. Census Bureau, U.S. Department of Commerce, 2002. (http://www.census.gov/prod/2002pubs/tp63rv.pdf)

Estimation of Median Incomes. The Census Bureau has changed the methodology for computing median income over time. The Census Bureau has computed medians using either Pareto interpolation or linear interpolation. Currently, we are using linear interpolation to estimate all medians. Pareto interpolation assumes a decreasing density of population within an income interval; whereas, linear interpolation assumes a constant density of population within an income interval. The Census Bureau calculated estimates of median income and associated standard errors for 1979 through 1987 using Pareto interpolation if the estimate was larger than $20,000 for people or $40,000 for families and households. This is because the width of the income interval containing the estimate is greater than $2,500.

We calculated estimates of median income and associated standard errors for 1976, 1977, and 1978 using Pareto interpolation if the estimate was larger than $12,000 for people or $18,000 for families and households. This is because the width of the income interval containing the estimate is greater than $1,000. All other estimates of median income and associated standard errors for 1976 through 2004 and almost all of the estimates of median income and associated standard errors for 1975 and earlier were calculated using linear interpolation.

Thus, use caution when comparing median incomes above $12,000 for people or $18,000 for families and households for different years. Median incomes below those levels are more comparable from year to year since they have always been calculated using linear interpolation. For an indication of the comparability of medians calculated using Pareto interpolation with medians calculated using linear interpolation, see Series P-60, No. 114, *Money Income in 1976 of Families and Persons in the United States.*

Standard Errors and Their Use. The sample estimate and its standard error enable one to construct a confidence interval. A confidence interval is a range that would include the average result of all possible samples with a known probability. For example, if all possible samples were surveyed under essentially the same general conditions and using the same sample design, and if an estimate and its standard error were calculated from each sample, then 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.

A particular confidence interval may or may not contain the average estimate derived from all possible samples. However, one can say with specified confidence that the interval includes the average estimate calculated from all possible samples.

Standard errors may be used to perform hypothesis testing. This is a procedure for distinguishing between population parameters using sample estimates. The most common type of hypothesis is that the population parameters are different. An example of this would be comparing the percentage of Whites in poverty to the percentage of Blacks in poverty.

Tests may be performed at various levels of significance. A significance level is the probability of concluding that the characteristics are different when, in fact, they are the same. For example, to conclude that two characteristics are different at the 0.10 level of significance, the absolute value of the estimated difference between characteristics must be greater than or equal to 1.645 times the standard error of the difference.

The Census Bureau uses 90-percent confidence intervals and 0.10 levels of significance to determine statistical validity. Consult standard statistical textbooks for alternative criteria.

Estimating Standard Errors. The Census Bureau uses replication methods to estimate the standard errors of CPS estimates. These methods primarily measure the magnitude of sampling error. However, they do measure some effects of nonsampling error as well. They do not measure systematic biases in the data due to nonsampling error. Bias is the average over all possible samples of the differences between the sample estimates and the true value.

Generalized Variance Parameters. It is possible to compute and present an estimate of the standard error based on the survey data for each estimate in a report, but there are a number of reasons why this is not done. A presentation of the individual standard errors would be of limited use, since one could not possibly predict all of the combinations of results that may be of interest to data users. Additionally, variance estimates are based on sample data and have variances of their own. Therefore, some method of stabilizing these estimates of variance, for example, by generalizing or averaging over time, may be used to improve their reliability.

Experience has shown that certain groups of estimates have a similar relationship between their variance and expected value. Modeling or generalization may provide more stable variance estimates by taking advantage of these similarities. The generalized variance function is a simple model that expresses the variance as a function of the expected value of the survey estimate. The parameters of the generalized variance function are estimated using direct replicate variances. These generalized variance parameters provide a relatively easy method to obtain approximate standard errors for numerous characteristics. In this source and accuracy statement, Table 3 provides the generalized variance parameters for labor force estimates, and Tables 4 and 5 provide generalized variance parameters for characteristics from the ASEC data. Table 6 contains the state factors and populations and Table 7 contains the regional factors and populations.

Standard Errors of Estimated Numbers. The approximate standard error, s_x, of an estimated number from this microdata file can be obtained using the formula:

$$s_x = \sqrt{ax^2 + bx} \tag{1}$$

where x is the size of the estimate and a and b are the parameters in Tables 3, 4, and 5 associated with the particular type of characteristic. When calculating standard errors from cross-tabulations involving different characteristics, use the set of parameters for the characteristic that will give the largest standard error.

For information on calculating standard errors for labor force data from the CPS which involve quarterly or yearly averages see "Explanatory Notes and Estimate of Error: Household Data" in *Employment and Earnings*, a monthly report published by the U.S. Bureau of Labor Statistics.

Illustration No. 1

Suppose there were 3,395,000 unemployed females in the civilian labor force. Use Formula (1) and the appropriate parameters from Table 3 to get

Illustration 1	
Number unemployed females in the civilian labor force (x)	3,395,000
a parameter (a)	-0.000031
b parameter (b)	2,782
Standard error	95,000
90% confidence interval	3,239,000 to 3,551,000

The standard error is calculated as

$$s_x = \sqrt{-0.000031 \times 3{,}395{,}000^2 + 2{,}782 \times 3{,}395{,}000} = 95{,}000$$

and the 90-percent confidence interval is calculated as 3,395,000 ± 1.645 × 95,000.

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 possible samples.

Illustration No. 2

Suppose that there were 13,027,000 children (under age18) in poverty. Use Formula (1) and the appropriate parameters from Table 4 to get

Illustration 2	
Number children in poverty (*x*)	13,027,000
a parameter (*a*)	-0.000050
b parameter (*b*)	4,072
Standard error	211,000
90% confidence interval	12,680,000 to 13,374,000

The standard error is calculated as

$$s_x = \sqrt{-0.000050 \times 13{,}027{,}000^2 + 4{,}072 \times 13{,}027{,}000} = 211{,}000$$

and the 90-percent confidence interval is calculated as 13,027,000 ± 1.645 × 211,000.

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 possible samples.

Standard Errors of Estimated Percentages. The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends on both the size of the percentage and its base. 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 parameter from Table 3, 4, or 5 as indicated by the numerator. However, for calculating standard errors for different characteristics of families in poverty, use the standard error of a ratio equation (see formula (8) in “Standard Errors of Ratios”).

The approximate standard error, $s_{x,p}$, of an estimated percentage can be obtained by using the formula:

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

Here x is the total number of people, families, households, or unrelated individuals in the base of the percentage, p is the percentage ($0 \le p \le 100$), and b is the parameter in Table 3, 4, or 5 associated with the characteristic in the numerator of the percentage.

Illustration No. 3
Suppose that there were 45,820,000 out of 291,155,000 people, or 15.7 percent, who did not have health insurance coverage. Use Formula (2) and the appropriate parameter from Table 4 to get

Illustration 3	
Percentage without health insurance coverage (*p*)	15.7
Base (*x*)	291,155,000
B parameter (*b*)	2,652
Standard error	0.11
90% confidence interval	15.5 to 15.9

The standard error is calculated as

$$s_{x,p} = \sqrt{\frac{2,652}{291,155,000} \times 15.7 \times (100 - 15.7)} = 0.11$$

The 90-percent confidence interval of the percentage of people without health insurance is calculated as 15.7 ± 1.645 × 0.11.

Standard Errors of Differences. The standard error of the difference between two sample estimates is approximately equal to

$$s_{x-y} = \sqrt{s_x^2 + s_y^2} \tag{3}$$

where s_x and s_y are the standard errors of the estimates, x and y. The estimates can be numbers, percentages, ratios, 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. However, if there is a high positive (negative) correlation between the two characteristics, the formula will overestimate (underestimate) the true standard error.

For information on calculating standard errors for labor force data from the CPS which involve differences in consecutive quarterly or yearly averages, consecutive month-to-month differences in estimates, and consecutive year-to-year differences in monthly estimates see “Explanatory Notes and Estimates of Error: Household Data” in *Employment and Earnings*, a monthly report published by the U.S. Bureau of Labor Statistics.

Illustration No. 4
Suppose there are 16,006,000 men aged 25 and over who are never married and 8,977,000 men aged 25 and over who are divorced. The apparent difference is 7,029,000. Use Formulas (1) and (3) and the appropriate parameters from Table 4 to get

Illustration 4			
	Never Married (x)	Divorced (y)	Difference
Number of males aged 25+	16,006,000	8,977,000	7,029,000
a parameter (a)	-0.000009	-0.000009	-
b parameter (b)	2,652	2,652	-
Standard error	200,000	152,000	251,000
90% confidence interval	15,677,000 to 16,335,000	8,727,000 to 9,227,000	6,616,000 to 7,442,000

The standard error of the difference is calculated as

$$s_{x-y} = \sqrt{200{,}000^2 + 152{,}000^2} = 251{,}000$$

and the 90-percent confidence interval around the difference is calculated as 7,029,000 ± 1.645 × 251,000. Since this interval does not include zero, we can conclude with 90 percent confidence that the number of never married men over age 24 was higher than the number of divorced men over age 24.

<u>Illustration No. 5</u>
Suppose the White poverty rate is 10.8 percent with a base of 233,702,000, and the Black poverty rate is 24.7 percent with a base of 36,423,000. The apparent difference is 13.9. Use Formulas (2) and (3) and the appropriate parameters from Table 4 to get

Illustration 5			
	White (x)	Black (y)	Difference
Poverty rate	10.8	24.7	13.9
Base (x)	233,702,000	36,423,000	-
b parameter (b)	5,282	5,282	-
Standard error	0.15	0.52	0.54
90% confidence interval	10.6 to 11.0	23.8 to 25.6	13.0 to 14.8

The standard error of the difference is calculated as

$$s_{x-y} = \sqrt{0.15^2 + 0.52^2} = 0.54$$

and the 90-percent confidence interval around the difference is calculated as 13.9 ± 1.645 × 0.54. Since this interval does not include zero, we can conclude with 90 percent confidence that the poverty rate for Blacks is higher than the poverty rate for Whites.

Standard Errors of Averages for Grouped Data. The formula used to estimate the standard error of an average for grouped data is

$$s_{\bar{x}} = \sqrt{\frac{b}{y}(S^2)} \qquad (4)$$

In this formula, *y* is the size of the base of the distribution and *b* is the parameter from Table 3, 4, or 5. The variance, S^2, is given by the following formula:

$$S^2 = \sum_{i=1}^{c} p_i \bar{x}_i^2 - \bar{x}^2 \qquad (5)$$

where $\bar{x}$, the average of the distribution, is estimated by

$$\bar{x} = \sum_{i=1}^{c} p_i \bar{x}_i \qquad (6)$$

c = the number of groups; *i* indicates a specific group, thus taking on values 1 through *c*.

p_i = estimated proportion of households, families or people whose values, for the characteristic (x-values) being considered, fall in group *i*.

$\bar{x}_i$ = $(Z_{i-1} + Z_i)/2$ where Z_{i-1} and Z_i are the lower and upper interval boundaries, respectively, for group *i*. $\bar{x}_i$ is assumed to be the most representative value for the characteristic for households, families, and unrelated individuals or people in group *i*. Group *c* is open-ended, i.e., no upper interval boundary exists. For this group the approximate average value is

$$\bar{x}_c = \frac{3}{2} Z_{c-1} \qquad (7)$$

Illustration No. 6

Suppose the average income deficit (the difference between the poverty threshold and actual income) for families in poverty is $7,775 with a variance of 6,477,000. Use the appropriate parameter from Table 4 and Formula (4) to get:

Illustration 6	
Average income deficit for families in poverty ($\bar{x}$)	$7,775
Variance (S^2)	6,477,000
Base (*y*)	7,854,000
b parameter (*b*)	5,282
Standard error	$66
90% confidence interval	$7,666 to $7,884

The standard error is calculated as

$$s_{\bar{x}} = \sqrt{\frac{5,282}{7,854,000}(6,477,000)} = 66$$

and the 90-percent confidence interval is calculated as \$7,775 ± 1.645 × \$66.

Standard Errors of Ratios. Certain estimates may be calculated as the ratio of two numbers. Compute the standard error of a ratio, *x*/*y*, using

$$s_{x/y} = \frac{x}{y}\sqrt{\left(\frac{s_x}{x}\right)^2 + \left(\frac{s_y}{y}\right)^2 - 2r\frac{s_x s_y}{xy}} \tag{8}$$

The standard error of the numerator, s_x, and that of the denominator, s_y, may be calculated using formulas described earlier. In Formula (8), *r* represents the correlation between the numerator and the denominator of the estimate.

For one type of ratio, the denominator is a count of families or households and the numerator is a count of people in those families or households with a certain characteristic. If there is at least one person with the characteristic in every family or household, use 0.7 as an estimate of *r*. An example of the type is the average number of children per family with children.

For all other types of ratios, *r* is assumed to be zero. If r is actually positive (negative), then this procedure will provide an overestimate (underestimate) of the standard error of the ratio. Examples of this type are the average number of children per family and the family poverty rate.

Note: For estimates expressed as the ratio of *x* per 100 *y* or *x* per 1,000 *y*, multiply Formula (8) by 100 or 1,000, respectively, to obtain the standard error.

Illustration No. 7

Suppose the number of males working part-time is 8,591,000, and the number of females working part-time is 17,122,000. The ratio of males working part-time to the number of females working part-time would be 0.502. Use Formulas (1) and (8) with *r* = 0 and the appropriate parameters from Table 3 to get

Illustration 7			
	Males (*x*)	Females (*y*)	Ratio
Number who work part-time	8,591,000	17,122,000	0.50
a parameter (*a*)	-0.000032	-0.000031	-
b parameter (*b*)	2,971	2,782	-
Standard error	152,000	196,000	0.011
90% confidence interval	8,341,000 to 8,841,000	16,800,000 to 17,444,000	0.48 to 0.52

The standard error is calculated as

$$s_{x/y} = \frac{8{,}591{,}000}{17{,}122{,}000}\sqrt{\left(\frac{152{,}000}{8{,}591{,}000}\right)^2 + \left(\frac{196{,}000}{17{,}122{,}000}\right)^2} = 0.011$$

and the 90-percent confidence interval is calculated as $0.50 \pm 1.645 \times 0.011$.

Standard Errors of Estimated Medians. The sampling variability of an estimated median depends on the form of the distribution and the size of the base. One can approximate the reliability of an estimated median by determining a confidence interval about it. (See "Standard Errors and Their Use" for a general discussion of confidence intervals.)

Estimate the 68-percent confidence limits of a median based on sample data using the following procedure.

1. Determine, using Formula (2), the standard error of the estimate of 50 percent from the distribution.

2. Add to and subtract from 50 percent the standard error determined in step 1. These two numbers are the percentage limits corresponding to the 68-percent confidence about the estimated median.

3. Using the distribution of the characteristic, determine upper and lower limits of the 68-percent confidence interval by calculating values corresponding to the two points established in step 2.

   Use the following formula to calculate the upper and lower limits.

$$X_{pN} = \frac{pN - N_1}{N_2 - N_1}(A_2 - A_1) + A_1 \tag{9}$$

where

X_{pN} = estimated upper and lower bounds for the confidence interval ($0 \le p \le 1$). For purposes of calculating the confidence interval, p takes on the values determined in step 2. Note that X_{pN} estimates the median when $p = 0.50$.

N = for distribution of numbers: the total number of units (people, households, etc.) for the characteristic in the distribution.

= for distribution of percentages: the value 100.

p = the values obtained in Step 2.

A_1, A_2 = the lower and upper bounds, respectively, of the interval containing X_{pN} .

N_1, N_2 = for distribution of numbers: the estimated number of units (people, households, etc.) with values of the characteristic greater than or equal to A_1 and A_2, respectively.

= for distribution of percentages: the estimated percentage of units (people, households, etc.) having values of the characteristic greater than or equal to A_1 and A_2, respectively.

4. Divide the difference between the two points determined in step 3 by two to obtain the standard error of the median.

Note: Median incomes and their standard errors calculated as below may differ from those in published tables showing income, since narrower income intervals were used in those calculations.

Illustration No. 8

Suppose you want to calculate the standard error of the median of total money income for families with the following distribution

Illustration 8			
Income Level	Number of Families	Cumulative Number of Families	Cumulative Percent of Families
Under $5,000	2,185,000	2,185,000	2.84
$5,000 to $9,999	2,072,000	4,257,000	5.53
$10,000 to $14,999	3,060,000	7,317,000	9.50
$15,000 to $24,999	8,241,000	15,558,000	20.20
$25,000 to $34,999	8,378,000	23,936,000	31.08
$35,000 to $49,999	11,407,000	35,343,000	45.89
$50,000 to $74,999	15,836,000	51,179,000	66.45
$75,000 to $99,999	10,338,000	61,517,000	79.87
$100,000 and over	15,502,000	77,019,000	100.00

1. Using Formula (2) with b = 1,249, the standard error of 50 percent on a base of 77,019,000 is about 0.20 percent.

2. To obtain a 68-percent confidence interval on an estimated median, add to and subtract from 50 percent the standard error found in step 1. This yields percentage limits of 49.80 and 50.20.

3. The lower and upper limits for the interval in which the percentage limits falls are $50,000 and $75,000, respectively.

Then, by addition, the estimated numbers of families with an income greater than or equal to $50,000 and $75,000 are 41,676,000 and 25,840,000, respectively.

Using Formula (9), the upper limit for the confidence interval of the median is found to be about

$$X_{pN} = \frac{0.4980 \times 77{,}019{,}000 - 41{,}676{,}000}{25{,}840{,}000 - 41{,}676{,}000}(75{,}000 - 50{,}000) + 50{,}000 = 55{,}242$$

Similarly, the lower limit is found to be about

$$X_{pN} = \frac{0.5020 \times 77{,}019{,}000 - 41{,}676{,}000}{25{,}840{,}000 - 41{,}676{,}000}(75{,}000 - 50{,}000) + 50{,}000 = 54{,}756$$

Thus, a 68-percent confidence interval for the median income for families is from $54,756 to $55,242.

4. The standard error of the median is, therefore,

$$\frac{55{,}242 - 54{,}756}{2} = 243$$

Standard Errors of Estimated Per Capita Deficits. Certain average values in reports associated with the ASEC data represent the per capita deficit for households of a certain class. The average per capita deficit is approximately equal to where

$$x = \frac{hm}{p} \tag{10}$$

h = number of households in the class
m = average deficit for households in the class
p = number of people in households in the class
x = average per capita deficit of people in households in the class.

To approximate standard errors for these averages, use the formula

$$s_x = \frac{hm}{p}\sqrt{\left(\frac{s_m}{m}\right)^2 + \left(\frac{s_p}{p}\right)^2 + \left(\frac{s_h}{h}\right)^2 - 2r\left(\frac{s_p}{p}\right)\left(\frac{s_h}{h}\right)} \tag{11}$$

In Formula (11), r represents the correlation between p and h.

For one type of average, the class represents households containing a fixed number of people. For example, h could be the number of three-person households. In this case, there is an exact correlation between the number of people in households and the number of households. Therefore, r = 1 for such households.

For other types of averages, the class represents households of other demographic types, for example, households in distinct regions, households in which the householder is of a certain age group, and owner-occupied and tenant-occupied households. In this and other cases in which the correlation between *p* and *h* is not perfect, use 0.7 as an estimate of *r*.

Illustration No. 9

Suppose there are 26,564,000 people living in families in poverty, and 7,854,000 families in poverty, with the average deficit income for families in poverty being $7,775 with a standard error of $66. Use Formulas (1), (10), and (11) and the appropriate parameters from Table 4 and *r* = 0.7 to get

Illustration 9				
	Number (*h*)	Number of people (*p*)	Average income deficit (*m*)	Average per capita deficit (*x*)
Value for families in poverty	7,854,000	26,564,000	$7,775	$2,299
a parameter (*a*)	+0.000052	-0.000018	-	-
b parameter (*b*)	1,243	5,282	-	-
Correlation (*r*)	-	-	-	0.7
Standard Error	114,000	357,000	$66	$32
90% confidence interval	7,666,000 to 8,042,000	25,977,000 to 27,151,000	$7,666 to $7,884	$2,246 to $2,352

The estimate of the average per capita deficit is calculated as

$$x = \frac{7{,}854{,}000 \times 7{,}775}{26{,}564{,}000} = 2{,}299$$

and the estimate of the standard error is calculated as

$$s_x = 2{,}299\sqrt{\left(\frac{66}{7{,}775}\right)^2 + \left(\frac{357{,}000}{26{,}564{,}000}\right)^2 + \left(\frac{114{,}000}{7{,}854{,}000}\right) + 2 \times 0.7 \times \left(\frac{357{,}000}{26{,}564{,}000}\right) \times \left(\frac{114{,}000}{7{,}584{,}000}\right)}$$

$$= 32$$

The 90-percent confidence interval is calculated as $2,299 ± 1.645 × $32.

Accuracy of State Estimates. The redesign of the CPS following the 1980 census provided an opportunity to increase efficiency and accuracy of state data. All strata are now defined within state boundaries. The sample is allocated among the states to produce state and national estimates with the required accuracy while keeping total sample size to a minimum. Improved accuracy of state data was achieved with about the same sample size as in the 1970 design.

Since the CPS is designed to produce both state and national estimates, the proportion of the total population sampled and the sampling rates differ among the states. In general, the smaller the population of the state the larger the sampling proportion. For example, in Vermont approximately 1 in every 250 households is sampled each month. In New York the sample is about 1 in every 2,000 households. Nevertheless, the size of the sample in New York is four times larger than in Vermont because New York has a larger population.

Standard Errors for State Estimates. The standard error for a state may be obtained by determining new state-level a and b parameters and then using these adjusted parameters in the standard error formulas mentioned previously. To determine a new state-level b parameter (b_{state}), multiply the b parameter from Table 3, 4, or 5 by the state factor from Table 6. To determine a new state-level a parameter (a_{state}), use the following.

(1) If the a parameter from Table 3, 4, or 5 is positive, multiply the a parameter by the state factor from Table 6.

(2) If the a parameter in Table 3, 4, or 5 is negative, calculate the new state-level a parameter as follows:

$$a_{state} = \frac{-b_{state}}{POP_{state}} \qquad (12)$$

where POP_{state} is the state population is found in Table 6.

Note: The Census Bureau recommends the use of three-year averages to compare estimates across states and two-year averages to evaluate changes in state estimates over time. See "Standard Errors of Data for Combined Years" and "Standard Errors of Two-Year Moving Averages."

<u>Illustration No. 10</u>

Suppose that the number of people living in New York who had completed a bachelor's degree or more is 4,082,000. Use Formulas (1) and (12) and the appropriate parameters, factors, and populations from Tables 4 and 6 to get

Illustration 10	
Number of people in NY with at least a bachelor's degree (*x*)	4,802,000
b parameter (*b*)	1,206
New York state factor	1.17
State population	18,959,323
State a parameter (a_{state})	-0.000074
State b parameter (b_{state})	1,411
Standard error	67,000

Obtain the state-level b parameter by multiplying the b parameter, 1,206, by the state factor, 1.17. This gives b_{state} = 1,206 × 1.17 = 1,411. Obtain the needed state-level a parameter by:

$$a_{state} = \frac{-1{,}411}{18{,}959{,}323} = -0.000074$$

The standard error of the estimate of the number of people in New York state who had completed a bachelor's degree or more can then be found by using Formula (1) and the new state-level a and b parameters, -0.000074 and 1,411, respectively. The standard error is given by:

$$s_x = \sqrt{-0.000074 \times 4{,}082{,}000^2 + 1{,}411 \times 4{,}802{,}000} = 67{,}000$$

Standard Errors of Regional Estimates. To compute standard errors for regional estimates, follow the steps for computing standard errors for state estimates found in "Standard Errors for State Estimates" using the regional factors and populations found in Table 7.

Standard Errors of Groups of States. The standard error calculation for a group of states is similar to the standard error calculation for a single state. First, calculate a new state group factor for the group of states. Then, determine new state group a and b parameters. Finally, use these adjusted parameters in the standard error formulas mentioned previously.

Use the following formula to determine a new state group factor:

$$state\_group\_factor = \frac{\sum_{i=1}^{n} POP_i \times state\_factor_i}{\sum_{i=1}^{n} POP_i} \tag{13}$$

where POP_i and $state\_factor_i$ are the population and factor for state i from Table 6.
To obtain a new state group b parameter (b_{state_group}), multiply the b parameter from Table 3, 4, or 5 by the state factor obtained by Formula (13). To determine a new state group a parameter (a_{state_group}), use the following.

(1) If the a parameter from Table 3, 4, or 5 is positive, multiply the a parameter by the state group factor determined by Formula (13).

(2) If the a parameter in Table 3, 4, or 5 is negative, calculate the new state group a parameter as follows:

$$a_{state\_group} = \frac{-b_{state\_group}}{\sum_{i=1}^{n} POP_i} \tag{14}$$

<u>Illustration No. 11</u>
Suppose the state group factor for the state group Illinois-Indiana-Michigan was required. The appropriate factor would be

$$state\_group\_factor = \frac{12{,}562{,}462 \times 1.13 + 6{,}170{,}284 \times 1.08 + 10{,}000{,}053 \times 1.09}{12{,}562{,}462 + 6{,}170{,}284 + 10{,}000{,}053} = 1.11$$

Standard Errors of Data for Combined Years. Sometimes estimates for multiple years are combined to improve precision. For example, suppose $\bar{x}$ is an average derived from n consecutive years' data, i.e., $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$, where the x_i are the standard error estimates for the individual years. Use the formulas described previously to estimate the standard error, s_x, of each year's estimate. Then the standard error of $\bar{x}$ is

$$s_{\bar{x}} = \frac{s_x}{n} \tag{15}$$

where

$$s_x = \sqrt{\sum_{i=1}^{n} s_{x_i}^2 + 2r \sum_{i=1}^{n-1} s_{x_i} s_{x_{i+1}}} \tag{16}$$

and s_{xi} are the standard errors of the estimates x_i over multiple years i. The correlation between consecutive years, r, is 0.30 for non-Hispanic people and 0.45 for Hispanic people. Correlation between nonconsecutive years is zero. The correlations were derived for income estimates but they can be used for other types of estimates where the year-to-year correlation between identical households is high. In published reports using the ASEC data, the Census Bureau uses three-year average estimates for state to state comparisons and also for certain race/ethnicity groups[4]. These reports use two-year average estimates to compare state and certain race estimate across years with a two-year moving average. See "Standard Errors of Two-Year Moving Averages."

<u>Illustration No. 12</u>
Supposed that the 2002-2004 three-year average percentage of people without health insurance in California is 18.4. The percentages and standard errors for 2002, 2003, and 2004 are 18.2, 18.4, and 18.7 percent and 0.43, 0.43, and 0.38, respectively. Use Formulas (15) and (16) and with r = 0.30 to get

Illustration 12				
	2002	2003	2004	2002-2004 avg
Percentage of people without health insurance in California (x)	18.2	18.4	18.7	18.4
Correlation (r)	-	-	-	0.30
Standard Error	0.43	0.43	0.37	0.28
90% confidence interval	18.1 to 19.3	17.7 to 19.1	17.5 to 18.9	17.9 to 18.9

[4] Estimates of characteristics of the American Indian and Alaska Native (AIAN) and Native Hawaiian and Other Pacific Islander (NHOPI) populations based on a single-year sample would be unreliable due to the small size of the sample that can be drawn from either population. Accordingly, such estimates are based on multiyear averages.

The standard error of the three-year average is calculated as

$$s_{\bar{x}} = \frac{0.84}{3} = 0.28$$

where

$$s_x = \sqrt{0.43^2 + 0.43^2 + 0.37^2 + (2 \times 0.30 \times 0.43 \times 0.43) + (2 \times 0.30 \times 0.43 \times 0.37)} = 0.84$$

The 90-percent confidence interval for the three-year percentage of people without health insurance in California is 18.4 ± 1.645 × 0.28.

Note: To calculate the standard errors of single year state estimates, see "Standard Errors of State Estimates."

Standard Errors of Two-Year Moving Averages. Two-year moving averages also improve precision for comparing across years by using two-year averages that overlap by a year. Use the formulas described previously to estimate the standard error, s_x, of each year's estimate. Then the standard error of the difference of the overlapping, or moving, averages is, $\bar{x}_{1,2} - \bar{x}_{2,3}$, is

$$s_{\bar{x}_{1,2} - \bar{x}_{2,3}} = \frac{1}{2}\sqrt{s_{x_1}^2 + s_{x_3}^2} \quad (17)$$

Illustration No. 13
Suppose that you want to calculate the standard error of the moving average of the poverty rate of American Indians/Alaska Natives (AIAN). If the average for 2002-2003 was 23.9 and the average for 2003-2004 was 24.4. The standard error for 2002 was 2.1 and the standard error for 2004 was 2.1. Use Formula (17) and these values to get

Illustration 13			
	2002, 2003 average	2003, 2004 average	avg(2002,2003)-avg(2003,2004)
Poverty rate of AIAN (*x*)	23.9	24.4	0.5
Standard error	2.07 (2002)	2.07 (2004)	1.46
90% confidence interval	-	-	-2.9 to 1.9

The standard error of the two-year moving average is calculated as

$$s_{\bar{x}_{1,2} - \bar{x}_{2,3}} = \frac{1}{2}\sqrt{2.07^2 + 2.07^2} = 1.46$$

and the 90-percent confidence interval around the difference of the moving averages is calculated as 0.5 ± 1.645 × 1.46. Since this interval includes zero, we cannot conclude with 90 percent confidence that the 2003-2004 average poverty rate of American Indians or Alaska

Natives was different than the 2002-2003 average poverty rate of American Indians or Alaska Natives.

Table 3. Parameters for Computation of Standard Errors for Labor Force Characteristics: March 2005		
Characteristic	a	b
Total or White		
Civilian Labor Force, Employed	-0.000016	3,068
Not in Labor Force	-0.000009	1,833
Unemployed	-0.000016	3,096
Civilian Labor Force, Employed, Not in Labor Force, and Unemployed		
Men	-0.000032	2,971
Women	-0.000031	2,782
Both sexes, 16 to 19 years	-0.000022	3,096
Black		
Civilian Labor Force, Employed, Not in Labor Force, and Unemployed	-0.000154	3,455
Men	-0.000336	3,357
Women	-0.000282	3,062
Both sexes, 16 to 19 years	-0.001531	3,455
Hispanic		
Civilian Labor Force, Employed, Not in Labor Force, and Unemployed	-0.000187	3,455
Men	-0.000363	3,357
Women	-0.000380	3,062
Both sexes, 16 to 19 years	-0.001822	3,455
API, AIAN, NH & OPI		
Civilian Labor Force, Employed, Not in Labor Force, and Unemployed	-0.000272	3,198
Men	-0.000569	3,198
Women	-0.000521	3,198
Both sexes, 16 to 19 years	-0.004146	3,198

NOTE: (1) These parameters are to be applied to basic CPS monthly labor force estimates.
(2) For foreign-born and noncitizen characteristics for Total and White, the a and b parameters should be multiplied by 1.3. No adjustment is necessary for foreign-born and noncitizen characteristics for Blacks, Hispanics, and APIs.
(3) API, AIAN, NH, and OPI are Asian and Pacific Islander, American Indian and Alaska Native, Native Hawaiian, and Other Pacific Islander, respectively.

Table 4. a and b Parameters for Standard Error Estimates for People and Families: 2004 ASEC								
Characteristics	Total or White		Black		API, AIAN, NH & OPI		Hispanic	
	a	b	a	b	a	b	a	b
PEOPLE								
Educational Attainment	-0.000005	1,206	-0.000032	1,364	-0.000087	1,364	-0.000028	922
Employment Characteristics	-0.000016	3,068	-0.000151	3,455	-0.000346	3,198	-0.000141	3,455
People by Family Income	-0.000011	2,494	-0.000067	2,855	-0.000183	2,855	-0.000086	2,855
Income	-0.000005	1,249	-0.000034	1,430	-0.000092	1,430	-0.000043	1,430
Health Insurance	-0.000009	2,652	-0.000067	3,809	-0.000188	3,809	-0.000091	3,809
Marital Status, Household and Family Characteristics								
Some household members	-0.000009	2,652	-0.000067	3,809	-0.000188	3,809	-0.000091	3,809
All household members	-0.000011	3,222	-0.000099	5,617	-0.000277	5,617	-0.000134	5,617
Mobility Characteristics (Movers)								
Educational Attainment, Labor Force, Marital Status, HH, Family, and Income	-0.000005	1,460	-0.000026	1,460	-0.000072	1,460	-0.000035	1,460
US, County, State, Region, or MSA	-0.000014	3,965	-0.000070	3,965	-0.000195	3,965	-0.000095	3,965
Below Poverty								
Total	-0.000018	5,282	-0.000093	5,282	-0.000260	5,282	-0.000126	5,282
Male	-0.000037	5,282	-0.000197	5,282	-0.000534	5,282	-0.000247	5,282
Female	-0.000036	5,282	-0.000176	5,282	-0.000507	5,282	-0.000259	5,282
Age								
Under 15	-0.000067	4,072	-0.000277	4,072	-0.000763	4,072	-0.000314	4,072
Under 18	-0.000050	4,072	-0.000214	4,072	-0.000621	4,072	-0.000261	4,072
15 and over	-0.000023	5,282	-0.000124	5,282	-0.000338	5,282	-0.000158	5,282
15 to 24	-0.000048	1,998	-0.000212	1,998	-0.000583	1,998	-0.000184	1,998
25 to 44	-0.000024	1,998	-0.000119	1,998	-0.000308	1,998	-0.000144	1,998
45 to 64	-0.000028	1,998	-0.000167	1,998	-0.000477	1,998	-0.000309	1,998
65 and over	-0.000057	1,998	-0.000449	1,998	-0.001320	1,998	-0.000910	1,998
Unemployment	-0.000016	3,096	-0.000151	3,455	-0.000346	3,198	-0.000141	3,455
FAMILIES, HOUSEHOLDS, OR UNRELATED INDIVIDUALS								
Income	-0.000005	1,140	-0.000029	1,245	-0.000080	1,245	-0.000037	1,245
Marital Status, HH and Family Characteristics, Educational Attainment, Population by Age/Sex	-0.000005	1,052	-0.000022	952	-0.000061	952	-0.000029	952
Poverty	+0.000052	1,243	+0.000052	1,243	+0.000052	1,243	+0.000052	1,243

NOTES:
(1) These parameters are to be applied to the 2005Annual Social and Economic Supplement data.
(2) API, AIAN, NH, and OPI are Asian and Pacific Islander, American Indian and Alaska Native, Native Hawaiian, and Other Pacific Islander, respectively.
(3) Hispanics may be of any race.
(4) The Total or White, Black, and API parameters are to be used for both “alone” and “in combination” race group estimates.
(5) For nonmetropolitan characteristics, multiply a and b parameters by 1.5. If the characteristic of interest in total state population, no subtotaled by race or ancestry, the a and b parameters are zero.
(6) For foreign-born and noncitizen characteristics for Total and White, the a and b parameters should be multiplied by 1.3. No adjustment is necessary for foreign-born and noncitizen characteristics for Blacks, APIs, and Hispanics.

Table 5. a and b Parameters for Standard Error Estimates for People and Families (Two or More Races): 2005 ASEC

Characteristics	Two or More	
	a	b
PEOPLE		
Educational Attainment	-0.000087	1,364
Employment Characteristics	-0.000151	3,455
People by Family Income	-0.000183	2,855
Income	-0.000092	1,430
Health Insurance	-0.000188	3,809
Marital Status, Household and Family Characteristics		
Some household members	-0.000188	3,809
All household members	-0.000277	5,617
Mobility Characteristics (Movers)		
Educational Attainment, Labor Force, Marital Status, HH, Family, and Income	-0.000072	1,460
US, County, State, Region, or MSA	-0.000195	3,965
Below Poverty		
Total	-0.000260	5,282
Male	-0.000534	5,282
Female	-0.000507	5,282
Age		
Under 15	-0.000763	4,072
Under 18	-0.000621	4,072
15 and over	-0.000338	5,282
15 to 24	-0.000583	1,998
25 to 44	-0.000308	1,998
45 to 64	-0.000477	1,998
65 and over	-0.001320	1,998
Unemployment	-0.000151	3,455
FAMILIES, HOUSEHOLDS, OR UNRELATED INDIVIDUALS		
Income	-0.000080	1,245
Marital Status, HH and Family Characteristics, Educational Attainment, Population by Age/Sex	-0.000061	952
Poverty	+0.000052	1,243

NOTES:
(1) These parameters are to be applied to the 2005 Annual Social and Economic Supplement data.
(2) Two or More Races refers to the group of cases self-classified as having two or more races.
(3) For nonmetropolitan characteristics, multiply a and b parameters by 1.5. If the characteristic of interest in total state population, no subtotaled by race or ancestry, the a and b parameters are zero.

Table 6. Factors for State Standard Errors and Parameters and State Populations: 2005

State	Factor	Population	State	Factor	Population
Alabama	1.05	4,466,174	Montana	0.24	916,118
Alaska	0.18	636,883	Nebraska	0.46	1,721,885
Arizona	1.23	5,761,249	Nevada	0.67	2,365,581
Arkansas	0.68	2,715,843	New Hampshire	0.34	1,292,238
California	1.25	35,631,764	New Jersey	1.12	8,623,446
Colorado	1.20	4,554,409	New Mexico	0.58	1,892,325
Connecticut	0.88	3,450,873	New York	1.17	18,959,323
Delaware	0.22	823,736	North Carolina	1.11	8,404,121
District of Columbia	0.18	537,389	North Dakota	0.16	618,710
Florida	1.12	17,346,628	Ohio	1.09	11,295,607
Georgia	1.08	8,710,318	Oklahoma	0.91	3,442,293
Hawaii	0.29	1,220,364	Oregon	1.01	3,569,000
Idaho	0.36	1,385,557	Pennsylvania	1.09	12,211,801
Illinois	1.13	12,562,462	Rhode Island	0.30	1,062,288
Indiana	1.08	6,170,284	South Carolina	1.06	4,130,837
Iowa	0.77	2,912,156	South Dakota	0.17	757,465
Kansas	0.73	2,680,682	Tennessee	1.08	5,770,033
Kentucky	1.05	4,079,404	Texas	1.28	22,259,461
Louisiana	1.05	4,418,278	Utah	0.44	2,387,483
Maine	0.39	1,304,185	Vermont	0.18	6160496
Maryland	1.13	5,493,445	Virginia	1.08	7,281,902
Massachusetts	1.06	6,327,181	Washington	1.15	6,143,200
Michigan	1.09	10,000,053	West Virginia	0.39	1,790,339
Minnesota	1.07	5,060,337	Wisconsin	1.10	5,448,669
Mississippi	0.71	2,842,620	Wyoming	0.15	500,516
Missouri	1.11	5,667,256			

NOTES: (1) The state population counts in this table are for the 0+ population.
(2) For foreign-born and noncitizen characteristics for Total and White, the a and b parameters should be multiplied by 1.3. No adjustment is necessary for foreign-born and noncitizen characteristics for Blacks, API, and Hispanics.

Table 7. Factors and Regional Standard Errors and Parameters and Regional Populations: 2005

Region	Factor	Population
Midwest	1.03	64,895,566
Northeast	1.05	53,847,831
South	1.08	104,578,501
West	1.10	66,964,449

NOTES: (1) The state population counts in this table are for the 0+ population.
(2) For foreign-born and noncitizen characteristics for Total and White, the a and b parameters should be multiplied by 1.3. No adjustment is necessary for foreign-born and noncitizen characteristics for Blacks, API, and Hispanics.