Source of the Data and Accuracy of the Estimates for the 2020 Annual Social and Economic Supplement Microdata File

SOURCE OF THE DATA

The data in this microdata file and the estimates in the reports *Income and Poverty in the United States: 2019*, *Health Insurance Coverage in the United States: 2019*, and *The Supplemental Poverty Measure: 2019* come from the 2020[1] Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS).[2] The U.S. Census Bureau conducts the CPS ASEC over a 3-month period in February, March, and April, with most of the data collection occurring in the month of March. The CPS ASEC uses two sets of questions, the basic CPS and a set of supplemental questions. The CPS, sponsored jointly by the 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 Census Bureau and the U.S. Bureau of Labor Statistics also jointly sponsor the CPS ASEC.

Basic CPS. The monthly CPS collects primarily labor force data about the civilian noninstitutionalized population living in the United States. The institutionalized population, which is excluded from the universe, consists primarily of the population in correctional institutions and nursing homes (98 percent of the 4.0 million institutionalized people in the 2010 Census). Starting in August 2017, college and university dormitories were also excluded from the universe because most of the residents had usual residences elsewhere. Interviewers ask questions concerning labor force participation of each member 15 years old and older in sample households. Typically, the week containing the nineteenth of the month is the interview week. The week containing the twelfth is the reference week (i.e., the week about which the labor force questions are asked).

The CPS uses a multistage probability sample based on the results of the decennial census, with coverage in all 50 states and the District of Columbia. The sample is continually updated to account for new residential construction. When files from the most recent decennial census become available, the Census Bureau gradually introduces a new sample design for the CPS.

Every ten years, the CPS first-stage sample is redesigned[3] reflecting changes based on the most recent decennial census. In the first stage of the sampling process, primary sampling units (PSUs)[4] were selected for sample. In the 2000 design, the United States was divided

---

1 For clarity and consistency throughout this report, the term "collection year" is the year the data is collected (in this case, 2020), and "data year" is the year about which the data are obtained (in this case, 2019). 2020 CPS ASEC asks questions of data year 2019, 2019 CPS ASEC asks questions of data year 2018, etc.

2 Portions of the health insurance data in the report are based on the American Community Survey (ACS). Please refer to the ACS Source and Accuracy Statement in U.S. Census Bureau (2019c).

3 For detailed information on the 2010 sample redesign, please see Bureau of Labor Statistics (2014).

4 The PSUs correspond to substate areas (i.e., counties or groups of counties) that are geographically contiguous.

into 2,025 PSUs. These were then grouped into 824 strata and one PSU was selected for sample from each stratum. In the 2010 sample design, the United States was divided into 1,987 PSUs. These PSUs were then grouped into 852 strata. Within each stratum, a single PSU was chosen for the sample, with its probability of selection proportional to its population as of the most recent decennial census. In the case of strata consisting of only one PSU, the PSU was chosen with certainty.

In April 2014, the Census Bureau began phasing out the 2000 sample and replaced it with the 2010 sample, creating a mixed sampling frame. Two simultaneous changes occurred during this phase-in period. First, within the PSUs selected for both the 2000 and 2010 designs, sample households from the 2010 design gradually replaced sample households from the 2000 design. Second, new PSUs selected for only the 2010 design gradually replaced outgoing PSUs selected for only the 2000 design. By July 2015, the new 2010 sample design was completely implemented and the sample came entirely from the 2010 redesigned sample.

Approximately 70,300 sampled addresses were selected from the sampling frame for the basic CPS. Based on eligibility criteria, ten percent of these sampled addresses were sent directly to computer-assisted telephone interviewing (CATI). The remaining sampled addresses were assigned to interviewers for computer-assisted personal interviewing (CAPI).[5] Of all addresses in sample, about 59,700 were determined to be eligible for interview. Interviewers obtained interviews at about 43,600 of the housing units at these addresses.[6] Noninterviews occur when the occupants are not found at home after repeated calls or are unavailable for some other reason. Table 1 summarizes historical changes in the CPS design.

The 2020 Annual Social and Economic Supplement. In addition to the basic CPS questions, interviewers asked supplementary questions for the CPS ASEC. They asked these questions of the civilian noninstitutionalized population and also of military personnel who live in households with at least one other civilian adult. The additional questions covered the following topics:

- Household and family characteristics.
- Marital status.
- Geographic mobility.

---

5 For further information on CATI and CAPI and the eligibility criteria, please see U.S. Census Bureau (2019e).

6 Due to government restrictions/health and safety concerns stemming from the spread of COVID-19, March CPS interviewing was impacted. Interviewing began Sunday, March 15th. On Friday, March 20th, personal visits with respondents were halted nationwide, resulting in telephone contacts only. Additionally, both CATI contact centers were closed as of Friday, March 20th. All cases remaining in CATI for ASEC follow-up were closed out and sent in to headquarters. Therefore, no CATI follow-up occurred after March 20th. These procedural changes resulted in higher nonresponse for both the basic CPS and the ASEC Supplement. For additional information on the impacts of COVID-19 on the CPS ASEC, please see Subsection “Impact of the Coronavirus Pandemic” within Section “Comparability of Data”.

- Foreign-born population.
- Income from the previous calendar year.
- Work status/occupation.
- Health insurance coverage.
- Program participation.
- Educational attainment.

Including the basic CPS sample, approximately 91,500 addresses were in sample for the CPS ASEC. About 79,400 sampled addresses were determined to be eligible for interview, and about 60,400 interviews were conducted (see Table 1).

The additional sample for the CPS ASEC provides more reliable data than the basic CPS 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 CPS ASEC, please refer to U.S. Census Bureau (2019e).

Table 1. Description of the March Basic Current Population Survey and Annual Social and Economic Supplement Sample Cases

Time period	Number of sample PSUs[A]	Basic CPS[B] sampled addresses eligible		Total (CPS ASEC[C]/ADS[D] + basic CPS) sampled addresses eligible	
		Interviewed	Not interviewed	Interviewed	Not interviewed
2020	852	43,600	16,100	60,400	19,000
2019	852	48,900	11,100	68,300	13,600
2018	852	50,800	9,900	67,900	11,500
2017	852	52,400	9,300	70,000	10,900
2016	852	52,000	9,100	69,500	10,600
2015	852	52,900	8,200	74,300	10,300
2014 Redesign[E]	824	17,200	2,200	22,700	2,600
2014 Traditional[F]	824	35,500	4,600	51,500	5,800
2014	824	52,700	6,800	--	--
2013	824	52,900	6,400	75,500	7,700
2012	824	53,300	5,800	75,100	7,200
2011	824	53,400	5,300	75,900	6,500
2010	824	54,100	4,600	77,000	5,700
2009	824	54,100	4,600	76,200	5,700
2008	824	53,800	5,100	75,900	6,400
2007	824	53,700	5,600	75,500	7,100
2006	824	54,000	5,400	76,000	7,100
2005	[G]754/824	54,400	5,700	76,500	7,500
2004	754	55,000	5,200	77,700	7,000
2003	754	55,500	4,500	78,300	6,800

Time period	Number of sample PSUs[A]	Basic CPS[B] sampled addresses eligible		Total (CPS ASEC[C]/ADS[D] + basic CPS) sampled addresses eligible	
		Interviewed	Not interviewed	Interviewed	Not interviewed
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	[H]629/729	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	[I]449/461	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

Source: U.S. Census Bureau, Current Population Survey, 1959-2020 Annual Social and Economic Supplement.

[A] PSUs are primary sampling units.

[B] CPS is the Current Population Survey.

[C] CPS ASEC is the Annual Social and Economic Supplement of the Current Population Survey.

[D] The CPS ASEC was referred to as the Annual Demographic Supplement (ADS) until 2002.

[E] The 2014 CPS ASEC Redesign indicates the subsample of the basic CPS households which received the redesigned ASEC questionnaire incorporating new income and health insurance questions.

[F] The 2014 CPS ASEC Traditional indicates the subsample of the basic CPS households which received the the same ASEC questionnaire that was used in the 2013 CPS ASEC.

[G] The Census Bureau redesigned the CPS following the Census 2000. During phase-in of the new design, addresses from the new and old designs were in the sample.

[H] The Census Bureau redesigned the CPS following the 1980 Decennial Census of Population and Housing.

[I] 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 controls of the civilian noninstitutionalized population of the United States, each state, and the District of

Columbia. These population controls[7] are prepared monthly as part of the Census Bureau's Population Estimates Program.

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

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

The population controls for the states are distributed by:

- Race (Black alone and all other race groups combined).
- Age (0-15, 16-44, and 45 and over).
- Sex.

The independent estimates by age, sex, 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 2010 Census data is updated using data on the components of population change (births, deaths, and net international migration) with net internal migration as an additional component in the state population controls.

The net international migration component of the population controls includes:

- Net international migration of the foreign born;
- Net migration between the United States and Puerto Rico;
- Net migration of natives to and from the United States; and
- Net movement of the Armed Forces population to and from the United States.

Because the latest available information on these components lags behind 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 CPS ASEC includes a further adjustment to give married and unmarried partners the same weight.

ACCURACY OF THE 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.

---

[7] For additional information on population controls, including details on the demographic characteristics used and net international components, please see Chapters 1-3 and Appendix: History of the Current Population Survey of U.S. Census Bureau (2019e).

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, the estimation of standard errors 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. There are several sources of nonsampling error that may occur during the development or execution of the survey. It can occur because of circumstances created by the interviewer, the respondent, the survey instrument, or the way the data are collected and processed. Some nonsampling errors, and examples of each, include:

- Measurement error: The interviewer records the wrong answer, the respondent provides incorrect information, the respondent estimates the requested. information, or an unclear survey question is misunderstood by the respondent.
- Coverage error: Some individuals who should have been included in the survey frame were missed.
- Nonresponse error: Responses are not collected from all those in the sample or the respondent is unwilling to provide information.
- Imputation error: Values are estimated imprecisely for missing data.
- Processing error: Forms may be lost, data may be incorrectly keyed, coded, or recoded, etc.

To minimize these errors, the Census Bureau applies quality control procedures during all stages of the production process including the design of the survey, the wording of questions, the review of the work of interviewers and coders, and the statistical review of reports.

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. For more details, please see Appendix C of U.S. Census Bureau (1993).

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 2020 ASEC, the basic CPS household-level unweighted nonresponse rate was 23.9 percent. The household-level unweighted nonresponse rate for the ASEC was an additional 19.7 percent. These two nonresponse rates lead to a combined supplement unweighted nonresponse rate of 38.9 percent.[8]

In accordance with Census Bureau and Office of Management and Budget Quality Standards, the Census Bureau will conduct an analysis to assess nonresponse bias in the 2020 CPS ASEC.

Responses are made up of complete interviews and sufficient partial interviews. A sufficient partial interview is an incomplete interview in which the household or person answered enough of the questionnaire for the supplement sponsor to consider the interview complete. The remaining supplement questions may have been edited or imputed to fill in missing values. Insufficient partial interviews are considered to be nonrespondents. Refer to the supplement overview attachment in the technical documentation for the specific questions deemed critical by the sponsor as necessary to answer in order to be considered a sufficient partial interview.

As a result of sufficient partial interviews being considered responses, individual items/questions have their own response and refusal rates. As part of the nonsampling error analysis, the item response rates, item refusal rates, and edits are reviewed. For the CPS ASEC, the unweighted item refusal rates range from 0.0 percent to 3.3 percent. The unweighted item allocation rates range from 23.3 percent to 74.1 percent.

Undercoverage. The concept of coverage with a survey sampling process is defined as the extent to which the total population that could be selected for sample "covers" the survey's target population. Missed housing units and missed people within sample households create undercoverage in the CPS. Overall CPS undercoverage for March 2020 is estimated to be about ten percent. CPS coverage varies with age, sex, and race. Generally, coverage is higher for females than for males and higher for non-Blacks than for Blacks. This differential coverage is a general problem for most household-based surveys.

The CPS weighting procedure mitigates bias from 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 origin, 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 poststratification divided by the independent population control. Table 2 shows March 2020 CPS coverage ratios by age and sex for certain race and Hispanic groups. The CPS coverage ratios can exhibit some variability from month to month.

---

[8] Because the ASEC is at the household level, the overall/combined ASEC response rate is a product of the basic CPS response rate and the ASEC response rate.

Table 2. Current Population Survey Coverage Ratios: March 2020

Age group	All people	Total		White alone		Black alone		Residual race[A]		Hispanic[B]	
		Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
0-15	0.85	0.85	0.86	0.89	0.90	0.72	0.69	0.80	0.83	0.77	0.78
16-19	0.83	0.84	0.83	0.87	0.86	0.69	0.68	0.80	0.84	0.81	0.81
20-24	0.75	0.76	0.74	0.80	0.76	0.63	0.63	0.68	0.76	0.80	0.73
25-34	0.79	0.76	0.82	0.81	0.86	0.51	0.67	0.79	0.76	0.69	0.76
35-44	0.88	0.86	0.91	0.90	0.94	0.67	0.80	0.81	0.83	0.75	0.85
45-54	0.90	0.89	0.92	0.90	0.95	0.80	0.81	0.88	0.90	0.81	0.91
55-64	0.98	0.97	0.98	0.99	1.01	0.84	0.92	0.94	0.86	0.89	0.90
65+	1.02	1.03	1.01	1.06	1.04	0.88	0.95	0.87	0.77	0.92	0.92
15+	0.90	0.89	0.91	0.92	0.94	0.71	0.80	0.83	0.81	0.79	0.84
0+	0.89	0.88	0.90	0.92	0.94	0.71	0.77	0.82	0.82	0.78	0.82

Source: U.S. Census Bureau, Current Population Survey, March 2020.

[A] The Residual race group includes cases indicating a single race other than White or Black, and cases indicating two or more races.

[B] Hispanics may be any race.

Note: For a more detailed discussion on the use of parameters for race and ethnicity, please see the "Generalized Variance Parameters" section.

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

Data users should be aware that estimates in the reports, *Income and Poverty in the United States: 2019*, *Health Insurance Coverage in the United States: 2019*, and *The Supplemental Poverty Measure: 2019,* use the internal CPS ASEC file. The Census Bureau must keep survey responses confidential, so disclosure avoidance techniques are applied to files prior to public release. Therefore, some estimates using the microdata files may differ from the estimates provided in the reports.

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

Caution should also be used when comparing CPS ASEC results from different years. Below, more detail is provided on several reasons for caution when comparing estimates across years.

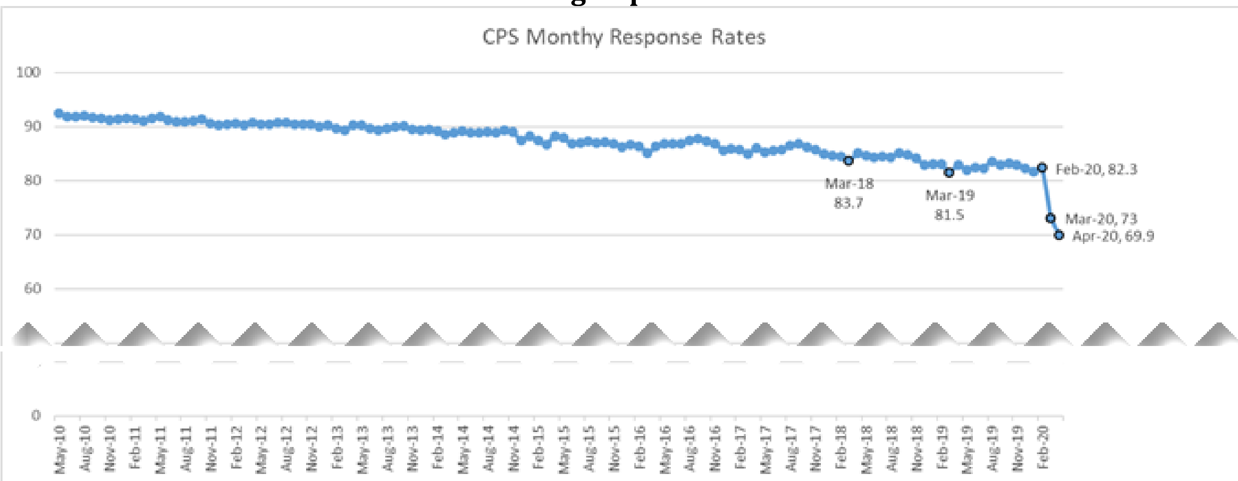
Impact of the Coronavirus Pandemic. Data users should exercise caution when comparing estimates for data year 2019 from the reports or from the microdata files to those from previous years due to the effects that the coronavirus (COVID-19) had on interviewing and response rates. Interviewing for the March CPS began on March 15th. In order to protect the health and safety of Census Bureau staff and respondents, the survey suspended in-person interviewing and closed the two CATI contact centers on March 20th. For the rest of March and through April, the Census Bureau continued to attempt all interviews by phone. For those whose first month in the survey was March or April, the Census Bureau used vendor-provided telephone numbers associated with the sample address.

While the Census Bureau went to great lengths to complete interviews by telephone, the response rate for the CPS basic household survey in March 2020 was 73[9] percent, about 10 percentage points lower than in preceding months and the same period in 2019. Further, as the Bureau of Labor Statistics (2020) stated in their Frequently Asked Questions accompanying the April 3rd release of The Employment Situation for March 2020, "Response rates for households normally more likely to be interviewed in person were particularly low. The response rate for households entering the sample for their first month was over 20 percentage points lower than in recent months, and the rate for those in the fifth month was over 10 percentage points lower."

The effect of changes in collection methods continued to be felt into April. The response rate for households entering the sample for their first month of interviewing was especially low. The unweighted April response rate for these households, which would normally have been interviewed in person, was over 30 percentage points lower than the average for the 12 months ending in February. Because the April ASEC selects households only from those in their first or fifth contact, the lower response rate translates into fewer potential ASEC households.

[9] This value differs from the response rate obtained using the values in the "Nonresponse" section because this value is specifically for March CPS whereas the values in the "Nonresponse" section are for the full CPS sample that was eligible for ASEC.

Figure 1: Unweighted Current Population Survey Monthly Response Rates for May 2010 through April 2020



Source: U.S. Census Bureau, Current Population Survey, internal data files, May 2010-April 2020.

The CPS ASEC response rate is complicated by the different months and samples that feed into the survey. Further, it includes an adjustment factor to account for those who responded to the basic survey but refused to answer the supplement. The Census Bureau estimates that the unweighted combined supplement response rate was 61.1 percent in 2020, down from 67.6 percent in 2019.

The change from conducting first interviews in person to making first contacts by telephone only is a contributing factor to the lower response rates. Further, it is likely that the characteristics of people for whom a telephone number was found may be systematically different from the people for whom the Census Bureau was unable to obtain a telephone number. While the Census Bureau creates weights designed to adjust for nonresponse and to control weighted counts to independent population estimates by age, sex, race, and Hispanic origin, the magnitude of the increase in (and differential nature of) nonresponse related to the pandemic likely reduced their effectiveness. Using administrative data, Census Bureau researchers have documented that there are more (and larger) differences between respondents and nonrespondents in 2020 than in the prior three years. Of particular interest for the estimates in the ASEC reports are the differences in median income and educational attainment, indicating that respondents in 2020 had relatively higher income and were more educated than nonrespondents.[10]

Change in Processing System. Data users should exercise caution when comparing estimates from the CPS ASEC for data years 2019 and 2018 to estimates from earlier years. An updated data processing system was implemented beginning with data year 2018 estimates. This system introduced demographic edit changes to account for same-sex couples, revised procedures for editing income and health insurance variables, and added several new income and health insurance variables. Changes to the editing procedures encompassed both changes to the resolution of logically inconsistent data and changes to the imputation methods. The

10 For additional information, please see Rothbaum & Bee (2020).

2019 and 2020 CPS ASEC estimates for data years 2018 and 2019 can be compared to the 2018 CPS ASEC Bridge Files[11], which contain data year 2017 estimates, and to the 2017 CPS ASEC Research Files[12], which contain estimates for data year 2016. The 2017 Research File and the 2018 Bridge File both use the new processing system and serve as a bridge between the legacy production files and the updated processing system. Data users should be aware that the estimates from the 2017 and 2018 CPS ASEC Files for data years 2016 and 2017 using the legacy processing system are not directly comparable to 2019 CPS ASEC and 2020 CPS ASEC estimates.

Change in Questionnaire. In 2014, the ASEC questionnaire was resigned to incorporate new income and health insurance questions. Due to the differences in measurement, health insurance estimates for 2014-2017 CPS ASEC for data years 2013-2016 are not directly comparable to health insurance estimates for previous years.[13] For income and poverty estimates, when survey changes had statistically significant impacts, comparisons should be made by adjusting historical published estimates to approximate the magnitude of those impacts.[14]

Change in Census-Based Controls. Data users should exercise caution when comparing estimates for 2019 from the microdata file or from the ASEC reports, *Income and Poverty in the United States: 2019* and *Health Insurance Coverage in the United States: 2019* (which reflect 2010 Census-based controls), with estimates from the microdata files or ASEC Reports for 2001 to 2010 (from March 2002 CPS to March 2011 CPS), which reflect 2000 Census-based controls, and to 1993 to 2000 (from March 1994 CPS to March 2001 CPS), which reflect 1990 Census-based controls. Ideally, the same population controls should be used when comparing any estimates. In reality, the use of the same population controls is not practical when comparing trend data over a period of 10 to 20 years. Thus, when it is necessary to combine or compare data based on different controls or different designs, data users should be aware that changes in weighting controls or weighting procedures could create small differences between estimates.

Microdata files from previous years reflect the latest available census-based controls. Although the most recent change in population controls had relatively little impact on summary measures such as averages, medians, and percentage distributions, it did have a significant impact on levels. For example, use of 2010 Census-based controls results in about a 0.2 percent increase from the 2000 Census-based controls in the civilian noninstitutionalized population and in the number of families and households. Thus, estimates of levels for data collected in 2012 and later years will differ from those for earlier years by more than what could be attributed to actual changes in the population.

---

11 For additional information on the 2018 CPS ASEC Bridge Files, please see the Documentation and User Notes in US Census Bureau (2019b).

12 For additional information on the 2017 CPS ASEC Research Files, please see the Documentation and User Notes in US Census Bureau (2019a).

13 For more information, see U.S. Census Bureau (2019f).

14 For more details on the adjustment for these comparisons, see U.S. Census Bureau (2019g).

These differences could be disproportionately greater for certain population subgroups than for the total population.

Users should also exercise caution because of changes caused by the phase-in of the 2010 Census files (see "Basic CPS").[15] During this time period, CPS data were collected from sample designs based on different censuses. Two features of the new CPS design have the potential of affecting estimates: (1) the temporary disruption of the rotation pattern from August 2014 through June 2015 for a comparatively small portion of the sample and (2) the change in sample areas. Most of the known effect on estimates during and after the sample redesign will be the result of changing from 2000 to 2010 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. The Census Bureau recommends that data users incorporate information about nonsampling errors into their analyses, as nonsampling error could impact the conclusions drawn from the results. 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 U.S. Census Bureau (2019e) and Brooks & Bailar (1978).

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. 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. All other estimates of median income and associated standard errors for 1976 through 2019 (2020 CPS ASEC), 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

[15] The phase-in process using the 2010 Census files began April 2014.

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 U.S. Census Bureau (1978) and U.S. Census Bureau (1993).

Standard Errors and Their Use. A sample estimate and its standard error enable one to construct a confidence interval. A confidence interval is a range about a given estimate that has a specified probability of containing the average result of all possible samples. 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, but one can say with the specified confidence that the interval includes the average estimate calculated from all possible samples.
Standard errors may also be used to perform hypothesis testing, 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 men who were part-time workers to the percentage of women who were part-time workers.

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.

The tables in *Income and Poverty in the United States: 2019, Health Insurance Coverage in the United States: 2019*, and *The Supplemental Poverty Measure: 2019* list estimates followed by a number labeled "Margin of Error (±)." This number can be added to and subtracted from the estimates to calculate upper and lower bounds of the 90-percent confidence interval. For example, *Health Insurance Coverage in the United States: 2019* shows the numbers for health insurance. For the statement, "8.0 percent of people were uninsured for the entire calendar year," the 90-percent confidence interval for the estimate, 8.0 percent, is 8.0 (± 0.2) percent, or 7.8 percent to 8.2 percent.[16]

[16] Note that the confidence interval here does not match the confidence interval given in Illustration 3 because the standard errors/margin of errors were calculated in two different ways. The margin of errors within the tables in the reports are calculated using direct estimates, whereas the standard errors within the illustrations later in this document are calculated using generalized variance estimates.

Estimating Standard Errors. The Census Bureau uses replication methods to estimate the standard errors of CPS and ASEC 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 associated with nonsampling error. Bias is the average over all possible samples of the differences between the sample estimates and the true value.

There are two ways to calculate standard errors for the 2020 CPS ASEC microdata file.

1. Direct estimates created from replicate weighting methods;
2. Generalized variance estimates created from generalized variance function (GVF) parameters *a* and *b*.

While replicate weighting methods provide the most accurate variance estimates, this approach requires more computing resources and more expertise on the part of the user. The GVF parameters provide a method of balancing accuracy with resource usage as well as a smoothing effect on standard error estimates. For more information on calculating direct estimates, refer to the "Replicate Weighting" section. For more information on GVF estimates, refer to the "Generalized Variance Parameters" section.

The *Income and Poverty in the United States: 2019*, *Health Insurance Coverage in the United States: 2019*, and *The Supplemental Poverty Measure: 2019* reports use replicate weights to calculate the margins of error of the estimates seen in tables and throughout the reports. In 2009, the Census Bureau released replicate weights for the 2005 through 2009 CPS ASEC collection years and has released replicate weights for each year since with the release of the CPS ASEC public use data. Since the published GVF parameters generally underestimated standard errors, standard errors produced using direct estimates may be higher than in previous reports. For most CPS ASEC estimates, the increase in standard errors from GVF to direct estimates will not alter the findings. However, marginally significant differences using the GVF may not be significant using replicate weights.

The examples in this source and accuracy statement are for guidance calculating standard errors using the generalized variance parameters. The use of generalized variance parameters is the recommended method of calculating standard errors for data users who do not have the ability to calculate the standard errors using replicate weights.

Replicate Weighting. The Census Bureau is releasing public use replicate weight files for the 2020 CPS ASEC that can be matched to the microdata files.

Replicate estimates are created using each of the 160 weights independently to create 160 replicate estimates. For point estimates, multiply the replicate weights by the item of interest at the record level (either an indicator variable to determine the number of people with a characteristic or a variable that contains some value) and tally the weighted values to create the 160 replicate estimates. Use these replicate estimates in formula (1) below to calculate the total variance for the item of interest. For example, say that the item of interest is the number of males. Tally the weights for all the records that indicated male to create the 160 replicate estimates of the number of males. Then use these estimates in the formula to calculate the total variance for the number of males.

Calculate variance estimates for the estimates using:

$$\mathrm{var}(\hat{\theta}_0) = \frac{4}{160}\sum_{i=1}^{160}(\hat{\theta}_i - \hat{\theta}_0)^2 \quad (1)$$

where $\hat{\theta}_0$ is the estimate of the statistic of interest, such as a point estimate or proportion, using the weight for the full sample, and $\hat{\theta}_i$ are the replicate estimates of the same statistic using the replicate weights. The standard error is the square root of the variance.
For more information on using replicate weights and calculating direct estimates, see U.S. Census Bureau (2009).

Generalized Variance Parameters. While it is possible to estimate the standard error based on the survey data for each estimate in a report, 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, data users have access to CPS microdata files, and it is impossible to compute in advance the standard error for every estimate one might obtain from those data sets. Moreover, variance estimates are based on sample data and have variances of their own. Therefore, some methods 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 similar relationships between their variances and expected values. Modeling or generalizing may provide more stable variance estimates by taking advantage of these similarities. The GVF is a simple model that expresses the variance as a function of the expected value of the survey estimate. The parameters of the GVF are estimated using direct replicate variances. These GVF parameters provide a relatively easy method to obtain approximate standard errors for numerous characteristics.

In this source and accuracy statement:

- Tables 4 through 17 provide illustrations for calculating standard errors;
- Table 18 provides the GVF parameters for labor force estimates;
- Table 19 provides GVF parameters for characteristics from the 2020 CPS ASEC;
- Tables 20 and 21 provide correlation coefficients for comparing estimates from consecutive years;
- Table 22 provides correlation coefficients between race and subgroups; and
- Tables 23 and 24 provide factors and population controls to derive state and regional parameters.

The basic CPS questionnaire records the race and ethnicity of each respondent. With respect to race, a respondent can be White, Black, Asian, American Indian and Alaskan Native (AIAN), Native Hawaiian and Other Pacific Islander (NHOPI), or combinations of two or more of the preceding. A respondent's ethnicity can be Hispanic or non-Hispanic, regardless of race.

The GVF parameters to use in computing standard errors are dependent upon the race/ethnicity group of interest. Table 3 summarizes the relationship between the race/ethnicity group of interest and the GVF parameters to use in standard error calculations.

Table 3. Estimation Groups of Interest and Generalized Variance Parameters

Race/ethnicity group of interest	Generalized variance parameters to use in standard error calculations
Total population	Total or White
White alone, White alone or in combination (AOIC), or White non-Hispanic population	Total or White
Black alone, Black AOIC, or Black non-Hispanic population	Black
Asian alone, Asian AOIC, or Asian non-Hispanic population	Asian, American Indian and Alaska Native (AIAN), Native Hawaiian and Other Pacific Islander (NHOPI)
AIAN alone, AIAN AOIC, or AIAN non-Hispanic population	Asian, AIAN, NHOPI
NHOPI alone, NHOPI AOIC, or NHOPI non-Hispanic population	Asian, AIAN, NHOPI
Populations from other race groups	Asian, AIAN, NHOPI
Hispanic[A] population	Hispanic[A]
Two or more races[B] – employment/unemployment and educational attainment characteristics	Black
Two or more races[B] – all other characteristics	Asian, AIAN, NHOPI

Source: U.S. Census Bureau, Current Population Survey, internal data files.

A Hispanics may be any race.

B Two or more races refers to the group of cases self-classified as having two or more races.

Note: The AOIC population for a race group of interest includes people reporting only the race group of interest (alone) and people reporting multiple race categories including the race group of interest (in combination).

When calculating standard errors for an estimate of interest from cross-tabulations involving different characteristics, use the set of GVF parameters for the characteristic that will give the largest standard error. If the estimate of interest is strictly from basic CPS data, the GVF parameters will come from the CPS GVF table (Table 18). If the estimate is using ASEC data, the GVF parameters will come from the ASEC GVF table (Table 19).

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

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

Here *x* is the size of the estimate, and *a* and *b* are the parameters in Table 18 or 19 associated with the particular type of characteristic.

<u>Illustration 1</u>
Suppose there were 3,826,000 unemployed females (ages 16 and up) in the civilian labor force. Table 4 shows how to use the appropriate parameters from Table 18 and Formula (2) to estimate the standard error and confidence interval.

Table 4. Illustration of Standard Errors of Estimated Numbers

Number of unemployed females in the civilian labor force (x)	3,826,000
a-parameter (a)	-0.000028
b-parameter (b)	2,788
Standard error	101,000
90-percent confidence interval	3,660,000 to 3,992,000

Source: U.S. Census Bureau, Current Population Survey, March 2020.

The standard error is calculated as

$$s_x = \sqrt{-0.000028 \times 3{,}826{,}000^2 + 2{,}788 \times 3{,}826{,}000},$$

which, rounded to the nearest thousand, is 101,000. The 90-percent confidence interval is calculated as 3,826,000 ± 1.645 × 101,000.

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.

<u>Illustration 2</u>
Suppose there were 62,342,000 married-couple family households. Table 5 shows how to use the appropriate parameters from Table 19 and Formula (2) to estimate the standard error and confidence interval.

Table 5. Second Illustration of Standard Errors of Estimated Numbers

Number of married-couple family households (x)	62,342,000
a-parameter (a)	-0.000009
b-parameter (b)	3,238
Standard error	409,000
90-percent confidence interval	61,669,000 to 63,015,000

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

The standard error is calculated as

$$s_x = \sqrt{-0.000009 \times 62{,}342{,}000^2 + 3{,}238 \times 62{,}342{,}000}$$

which, rounded to the nearest thousand, is 409,000. The 90-percent confidence interval is calculated as 62,342,000 ± 1.645 × 409,000.

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 18 or 19 as indicated by the numerator.

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

$$s_{y,p} = \sqrt{\frac{b}{y} p(100 - p)} \qquad (3)$$

Here y is the total number of people, families, households, or unrelated individuals in the base or denominator of the percentage, p is the percentage $100*x/y$ ($0 \le p \le 100$), and b is the parameter in Table 18 or 19 associated with the characteristic in the numerator of the percentage.

Illustration 3

The report, *Health Insurance Coverage in the United States: 2019*, shows that there were 26,111,000 out of 324,550,000 people, or 8.0 percent, who did not have health insurance. Table 6 shows how to use the appropriate parameters from Table 19 and Formula (3) to estimate the standard error and confidence interval.

Table 6. Illustration of Standard Errors of Estimated Percentages

Percentage of people without health insurance (p)	8.0
Base (y)	324,550,000
b-parameter (b)	3,022
Standard error	0.08
90-percent confidence interval	7.9 to 8.1

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

The standard error is calculated as

$$s_{y,p} = \sqrt{\frac{3{,}022}{324{,}550{,}000} \times 8.0 \times (100.0 - 8.0)} = 0.08$$

and the 90-percent confidence interval for the estimated percentage of people without health insurance is from 7.9 to 8.1 percent (i.e., 8.0 ± 1.645 × 0.08).

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

$$s_{x_1-x_2} = \sqrt{{s_{x_1}}^2 + {s_{x_2}}^2 - 2rs_{x_1}s_{x_2}} \qquad (4)$$

where s_{x_1} and s_{x_2} are the standard errors of the estimates, x_1 and x_2. The estimates can be numbers, percentages, ratios, etc. Tables 20 and 21 contain the correlation coefficient, *r*, for CPS year-to-year comparisons for CPS poverty, income, and health insurance estimates of numbers and proportions. Table 22 contains the correlation coefficient *r* for making comparisons between race categories that are subsets of one another. For example, to compare the number of people in poverty who listed White as their only race to the number of people in poverty who are White alone or in combination with another race, a correlation coefficient is needed to account for the large overlap between the two groups. For making other comparisons (including race overlapping where one group is not a complete subset of the other), assume that *r* equals zero. Making this assumption will result in accurate estimates of standard errors 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.

Illustration 4

Suppose there were 25,886,000 men over age 24 who were never married and 10,626,000 men over age 24 who were divorced. The apparent difference is 15,260,000. Table 7 shows how to use Formulas (2) and (4) with *r* = 0 and the appropriate parameters from Table 19 to estimate the standard errors and confidence intervals.

Table 7. Illustration of Standard Errors of Estimated Differences

	Never married (x_1)	Divorced (x_2)	Difference
Number of males over age 24	25,886,000	10,626,000	15,260,000
a-parameter (a)	-0.000009	-0.000009	-
b-parameter (b)	2,808	2,808	-
Standard error	258,000	170,000	309,000
90-percent confidence interval	25,462,000 to 26,310,000	10,346,000 to 10,906,000	14,752,000 to 15,768,000

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

The standard error of the difference is calculated as

$$s_{x_1-x_2} = \sqrt{258{,}000^2 + 170{,}000^2}$$

which, rounded to the nearest thousand, is 309,000. The 90-percent confidence interval around the difference is calculated as 15,260,000 ± 1.645 × 309,000. 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.

Illustration 5

The report, *Income and Poverty in the United States: 2019*, shows that 11,869,000 out of 73,284,000 children, or 16.2 percent, were reported as in poverty in 2018, and that 10,466,000 out of 72,637,000, or 14.4 percent, were in poverty in 2019. The apparent difference is 1.8 percent. Table 8 shows how to use the appropriate parameters from Table 19 and Formulas (3) and (4) to estimate the standard error and confidence interval.

Table 8. Illustration of Standard Errors of Estimated Differences

	2018 (x_1)	2019 (x_2)	Difference
Percentage of children in poverty (p)	16.2	14.4	1.8
Base	73,284,000	72,637,000	-
b-parameter (b)	2,718[A]	3,781	-
Correlation coefficient (r)	-	-	0.45
Standard error	0.22	0.25	0.25
90-percent confidence interval	15.8 to 16.6	14.0 to 14.8	1.4 to 2.2

Source: U.S. Census Bureau, Current Population Survey, 2019-2020 Annual Social and Economic Supplement.

[A] This value comes from the Source and Accuracy Statement for the 2019 Annual Social and Economic Supplement, Appendix G, Table 19 in U.S. Census Bureau (2019d). For additional information, see the "Year-to-Year Factors" section.

The standard error of the difference is calculated as

$$s_{x_1-x_2} = \sqrt{0.22^2 + 0.25^2 - 2 \times 0.45 \times 0.22 \times 0.25} = 0.25$$

and the 90-percent confidence interval around the difference is calculated as 1.8 ± 1.645 × 0.25. Since this interval does not include zero, we can conclude with 90-percent confidence that the percentage of children in poverty in 2019 is significantly less than the percentage of children in poverty in 2018.

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

$$s_{x/y} = \frac{x}{y}\sqrt{\left(\frac{s_x}{x}\right)^2 + \left(\frac{s_y}{y}\right)^2 - 2r\frac{s_x s_y}{xy}} \qquad (5)$$

The standard error of the numerator, s_x, and that of the denominator, s_y, may be calculated using formulas described earlier. In Formula (5), 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 this type is the average number of children per family with children.

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

Note: For estimates expressed as the ratio of *x* per 100 *y* or *x* per 1,000 *y*, multiply Formula (5) by 100 or 1,000, respectively, to obtain the standard error.

Illustration 6
Suppose there were 11,328,000 males working part-time and 17,534,000 females working part-time. The ratio of males working part-time to females working part-time would be 0.646, or 64.6 percent. Table 9 shows how to use the appropriate parameters from Table 18 and Formulas (2) and (5) with $r = 0$ to estimate the standard errors and confidence intervals.

Table 9. Illustration of Standard Errors of Estimated Ratios

	Males (x)	Females (y)	Ratio
Number who work part-time	11,328,000	17,534,000	0.646
a-parameter (a)	-0.000031	-0.000028	-
b-parameter (b)	2,947	2,788	-
Standard error	171,000	201,000	0.012
90-percent confidence interval	11,047,000 to 11,609,000	17,203,000 to 17,865,000	0.626 to 0.666

Source: U.S. Census Bureau, Current Population Survey, March 2020.

The standard error is calculated as

$$s_{x/y} = \frac{11{,}328{,}000}{17{,}534{,}000}\sqrt{\left(\frac{171{,}000}{11{,}328{,}000}\right)^2 + \left(\frac{201{,}000}{17{,}534{,}000}\right)^2} = 0.012$$

and the 90-percent confidence interval is calculated as $0.646 \pm 1.645 \times 0.012$.

Illustration 7
The report, *Income and Poverty in the United States: 2019*, shows that the number of families below the poverty level, x, was 6,554,000 and the total number of families, y, was 83,698,000. The ratio of families below the poverty level to the total number of families would be 0.078 or 7.8 percent. Table 10 shows how to use the appropriate parameters from Table 19 and Formulas (2) and (5) with $r = 0$ to estimate the standard errors and confidence intervals.

Table 10. Second Illustration of Standard Errors of Estimated Ratios

	In poverty (x)	Total (y)	Ratio (in percent)
Number of families	6,554,000	83,698,000	7.8
a-parameter (a)	0.000103	-0.000009	-
b-parameter (b)	5,529	3,238	-
Standard error	202,000	456,000	0.24
90-percent confidence interval	6,222,000 to 6,886,000	82,948,000 to 84,448,000	7.4 to 8.2

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

The standard error is calculated as

$$s_{x/y} = \frac{6{,}554{,}000}{83{,}698{,}000}\sqrt{\left(\frac{202{,}000}{6{,}554{,}000}\right)^2 + \left(\frac{456{,}000}{83{,}698{,}000}\right)^2} = 0.0024 = 0.24\%$$

and the 90-percent confidence interval of the percentage is calculated as 7.8 ± 1.645 × 0.24.

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. Using Formula (3) and the base of the distribution, calculate the standard error of 50 percent.

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 interval 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.

   Note: The percentage limits found in step 2 may or may not fall in the same characteristic distribution interval.

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

$$X_p = \frac{pN - N_1}{N_2 - N_1}(A_2 - A_1) + A_1 \qquad (6)$$

   where

X_p = 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_p 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_p.

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

4. Divide the difference between the two points determined in step 3 by 2 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 and reports showing income, since narrower income intervals were used in those calculations.

Illustration 8

The report, *Income and Poverty in the United States: 2019*, shows that there were 128,451,000 households, and their income was distributed as shown in Table 11.

Table 11. Distribution of Household Income for Illustration 8

Income level	Number of households	Cumulative number of households	Cumulative percent of households
Under $5,000	3,821,000	3,821,000	2.97%
$5,000 to $9,999	2,833,000	6,654,000	5.18%
$10,000 to $14,999	5,003,000	11,657,000	9.08%
$15,000 to $24,999	10,287,000	21,944,000	17.08%
$25,000 to $34,999	10,828,000	32,772,000	25.51%
$35,000 to $49,999	14,980,000	47,752,000	37.18%
$50,000 to $74,999	21,057,000	68,809,000	53.57%
$75,000 to $99,999	15,923,000	84,732,000	65.96%
$100,000 and over	43,719,000	128,451,000*	100.00%*

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.
*There may be a difference due to rounding.

1. Using Formula (3) with b = 3,938, the standard error of 50 percent on a base of 128,451,000 is about 0.28 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 49.72 and 50.28.

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

   Then the estimated numbers of households with an income less than or equal to $50,000 and $75,000 are 47,752,000 and 68,809,000, respectively.

   Using Formula (6), the lower limit for the confidence interval of the median is found to be about

$$X_{0.4972} = \frac{0.4972 \times 128{,}451{,}000 - 47{,}752{,}000}{68{,}809{,}000 - 47{,}752{,}000}(75{,}000 - 50{,}000) + 50{,}000 = 69{,}131$$

   Similarly, the upper limit is found to be about

$$X_{0.5028} = \frac{0.5028 \times 128{,}451{,}000 - 47{,}752{,}000}{68{,}809{,}000 - 47{,}752{,}000}(75{,}000 - 50{,}000) + 50{,}000 = 69{,}985$$

   Thus, a 68-percent confidence interval for the median income for households is from $69,131 to $69,985.

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

$$\frac{69{,}985 - 69{,}131}{2} = 427.0$$

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

$$s_{\bar{x}} = \sqrt{\frac{b}{y}(S^2)} \quad (7)$$

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

$$S^2 = \sum_{i=1}^{c} p_i \bar{x}_i^2 - \bar{x}^2 \quad (8)$$

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

$$\bar{x} = \sum_{i=1}^{c} p_i \bar{x}_i \quad (9)$$

where

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 being considered fall in group *i*.

$\bar{x}_i$ = $(Z_{Li} + Z_{Ui})/2$ where Z_{Li} and Z_{Ui} 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 of households, families, or people in group *i*. If group *c* is open-ended, i.e., no upper interval boundary exists, use a group approximate average value of

$$\bar{x}_c = \frac{3}{2} Z_{L_c} \quad (10)$$

<u>Illustration 9</u>
The report, *Income and Poverty in the United States: 2019*, shows that there were 6,554,000 families in poverty. Table 12 shows the distribution of the income deficit (the difference between their family income and poverty threshold) for all families in poverty.

Table 12. Distribution of Income Deficit for Illustration 9

Income deficit	Number of families in poverty	Percentage of families in poverty (p_i)	Average income deficit ($\overline{x}_i$)
Under $1000	468,000	7.1%	500
$1000 to $2,499	514,000	7.8%	1,750
$2,500 to $4,999	899,000	13.7%	3,750
$5,000 to $7,499	805,000	12.3%	6,250
$7,500 to $9,999	760,000	11.6%	8,750
$10,000 to $12,499	589,000	9.0%	11,250
$12,500 to $14,999	528,000	8.1%	13,750
$15,000 and over	1,991,000	30.4%	22,500
Total	6,554,000*	100%*	

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.
*There may be a difference due to rounding.

Using Formula (9),

$$\bar{x} = (0.071 \times 500) + (0.078 \times 1{,}750) + (0.137 \times 3{,}750) + (0.123 \times 6{,}250) + (0.116 \times 8{,}750) + (0.090 \times 11{,}250) + (0.081 \times 13{,}750) + (0.304 \times 22{,}500) = 11{,}436$$

and Formula (8),

$$S^2 = (0.071 \times 500^2) + (0.078 \times 1{,}750^2) + (0.137 \times 3{,}750^2) + (0.123 \times 6{,}250^2) + (0.116 \times 8{,}750^2) + (0.090 \times 11{,}250^2) + (0.081 \times 13{,}750^2) + (0.304 \times 22{,}500^2) - 11{,}436^2 = 65{,}692{,}000$$

Table 13 shows how to use the appropriate parameter from Table 19 and Formula (7) to estimate the standard error and confidence interval.

Table 13. Illustration of Standard Errors of Averages for Grouped Data

Average income deficit for families in poverty ($\bar{x}$)	$11,436
Variance (S^2)	65,692,000
Base (y)	6,554,000
b-parameter (b)	5,529
Standard error	$235
90-percent confidence interval	$11,049 to $11,823

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

The standard error is calculated as

$$s_{\bar{x}} = \sqrt{\frac{5{,}529}{6{,}554{,}000}(65{,}692{,}000)} = 235$$

and the 90-percent confidence interval is calculated as $11,436 ± 1.645 × $235.

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

$$x = \frac{hm}{p} \tag{11}$$

where

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

$$s_x = \frac{hm}{p}\sqrt{\left(\frac{s_m}{m}\right)^2 + \left(\frac{s_p}{p}\right)^2 + \left(\frac{s_h}{h}\right)^2 - 2r\left(\frac{s_p}{p}\right)\left(\frac{s_h}{h}\right)} \tag{12}$$

In Formula (12), 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 3-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 10

The report, *Income and Poverty in the United States: 2019*, shows that there were 22,431,000 people living in families in poverty, and 6,554,000 families in poverty, with an average deficit income for families in poverty of $11,436 with a standard error of $235 (from Illustration 9). Table 14 shows how to use Formulas (2), (11), and (12) and the appropriate parameters from Table 19 and $r = 0.7$ to estimate the standard errors and confidence intervals.

Table 14. Illustration of Standard Errors of Estimated Medians

	Number (h)	Number of people (p)	Average income deficit (m)	Average per capita deficit (x)
Value for families in poverty	6,554,000	22,431,000	$11,436	$3,341
a-parameter (a)	0.000103	-0.000113	-	-
b-parameter (b)	5,529	3,838	-	-
Correlation (r)	-	-	-	0.7
Standard error	202,000	171,000	$235	$111
90-percent confidence interval	6,222,000 to 6,886,000	22,150,000 to 22,712,000	$11,049 to $11,823	$3,158 to $3,524

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

The estimate of the average per capita deficit is calculated as

$$x = \frac{6{,}554{,}000 \times 11{,}436}{22{,}431{,}000} = 3{,}341$$

and the standard error is calculated as

$$s_x = \frac{6{,}554{,}000 \times 11{,}436}{22{,}431{,}000}\sqrt{\left(\frac{235}{11{,}436}\right)^2 + \left(\frac{171{,}000}{22{,}431{,}000}\right)^2 + \left(\frac{202{,}000}{6{,}554{,}000}\right)^2 - 2 \times 0.7 \times \left(\frac{171{,}000}{22{,}431{,}000}\right) \times \left(\frac{202{,}000}{6{,}554{,}000}\right)}$$
$$= 111$$

The 90-percent confidence interval is calculated as $3,341 ± 1.645 × $111.

<u>Accuracy of State Estimates</u>. 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.

Note: The Census Bureau recommends the use of 3-year averages to compare estimates across states and 2-year averages to evaluate changes in state income and poverty estimates over time. See "Standard Errors of Data for Combined Years." Further, the *Income and Poverty in the United States* report no longer presents state estimates. Therefore, the Census Bureau recommends the American Community Survey (ACS) microdata file as the preferred source for income and poverty state data in years 2006 (2005 estimates) to the present. A questionnaire redesign introduced with the 2014 CPS ASEC and an updated processing system introduced with the 2019 CPS ASEC each mark the start of new time series for health insurance estimates in the CPS ASEC, so data users should not create multiyear averages across these years.

Standard Errors of 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 18 or 19 by the state factor from Table 23. To determine a new state-level a-parameter (a_{state}), use the following:

(1) If the a-parameter from Table 18 or 19 is positive, multiply it by the state factor from Table 23.

(2) If the a-parameter in Table 18 or 19 is negative, calculate the new state-level a-parameter as follows:

$$a_{state} = \frac{-b_{state}}{POP_{state}} \quad (13)$$

where POP_{state} is the state population found in Table 23.

Illustration 11

Suppose there were 14,201,000 people living in New York state who were born in the United States. Table 15 shows how to use Formulas (2) and (13) and the appropriate parameter, factor, and population from Tables 19 and 23 to estimate the standard error and confidence interval.

Table 15. Illustration of Standard Errors of State Estimates

Number of people in New York born in the U.S. (x)	14,201,000
b-parameter (b)	2,808
New York state factor	1.19
State population	19,173,378
State b-parameter (b_{state})	3,342
State a-parameter (a_{state})	-0.000174
Standard error	111,000
90-percent confidence interval	14,018,000 to 14,384,000

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

Obtain the state-level b-parameter by multiplying the b-parameter, 2,808 by the state factor, 1.19. This gives b_{state} = 2,808 × 1.19 = 3,342. Obtain the needed state-level a-parameter by

$$a_{state} = \frac{-3{,}342}{19{,}173{,}378} = -0.000174$$

The standard error of the estimate of the number of people in New York state who were born in the United States can then be found by using Formula (2) and the new state-level *a*- and *b*- parameters, -0.000174 and 3,342, respectively. The standard error is given by

$$s_x = \sqrt{-0.000174 \times 14{,}201{,}000^2 + 3{,}342 \times 14{,}201{,}000}$$

which, rounded to the nearest thousand, is 111,000.

<u>Standard Errors of Regional Estimates</u>. 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 24.

<u>Illustration 12</u>
The report, *Income and Poverty in the United States: 2019*, shows that there were 14,845,000 of 124,032,005 people, or 12.0 percent, living in poverty in the South. Table 16 shows how to use Formulas (3) and (13) and the appropriate parameter, factor, and population from Tables 19 and 24 to estimate the standard error and confidence interval.

Table 16. Illustration of Standard Errors of Regional Estimates

Poverty rate in the South (p)	12.0
Base (y)	124,032,005
b-parameter (b)	3,838
South regional factor	1.13
Regional b-parameter (b_{region})	4,337
Standard error	0.19
90-percent confidence interval	11.7 to 12.3

Source: U.S. Census Bureau, Current Population Survey, 2020 Annual Social and Economic Supplement.

Obtain the region-level b-parameter by multiplying the b-parameter, 3,838, by the South regional factor, 1.13. This gives b_{region} = 3,838 × 1.13 = 4,337

The standard error of the estimate of the poverty rate for people living in the South can then be found by using Formula (3) and the new region-level b-parameter, 4,337. The standard error is given by

$$s_{y,p} = \sqrt{\frac{4{,}337}{124{,}032{,}005} \times 12.0 \times (100 - 12.0)} = 0.19$$

and the 90-percent confidence interval of the poverty rate for people living in the South is calculated as 12.0 ± 1.645 × 0.19.

<u>Standard Errors of Groups of States</u>. 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:

$$state\ group\ factor = \frac{\sum_{i=1}^{n} POP_i \times state\ factor_i}{\sum_{i=1}^{n} POP_i} \tag{14}$$

where POP_i and $state\ factor_i$ are the population and factor for state *i* from Table 23. To obtain a new state group b-parameter ($b_{state\ group}$), multiply the b-parameter from Table 18 or 19 by the state group factor obtained by Formula (14). To determine a new state group a-parameter ($a_{state\ group}$), use the following:

(1) If the a-parameter from Table 18 or 19 is positive, multiply it by the state group factor determined by Formula (14).

(2) If the a-parameter in Table 18 or 19 is negative, calculate the new state group a-parameter as follows:

$$a_{state\ group} = \frac{-b_{state\ group}}{\sum_{i=1}^{n} POP_i} \tag{15}$$

Illustration 13

Suppose the state group factor for the state group Illinois-Indiana-Michigan was required. The appropriate factor would be

$$state\ group\ factor = \frac{12{,}451{,}406 \times 1.17 + 6{,}657{,}419 \times 1.11 + 9{,}883{,}888 \times 1.11}{12{,}451{,}406 + 6{,}657{,}419 + 9{,}883{,}888} = 1.14$$

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 estimates for the individual years. Use the formulas described previously to estimate the standard error, s_{x_i}, of each year's estimate. Then the standard error of $\bar{x}$ is

$$s_{\bar{x}} = \frac{s_x}{n} \tag{16}$$

where

$$s_x = \sqrt{\sum_{i=1}^{n} s_{x_i}^2 + 2r \sum_{i=1}^{n-1} s_{x_i} s_{x_{i+1}}} \tag{17}$$

and s_{x_i} are the standard errors of the estimates x_i. Tables 20 and 21 contain the correlation coefficients, r, for the correlation between consecutive years i and i+1. Correlation between nonconsecutive years is zero. The correlations were derived for income, poverty, and health insurance estimates, but they can be used for other types of estimates where the year-to-year correlation between identical households is high.

The Census Bureau recommends the use of 3-year average estimates for certain small population subgroups[17] (see also "Accuracy of State Estimates.") Two-year moving averages are recommended for these small population subgroups for comparisons across adjacent years.

Illustration 14

The report, *Income and Poverty in the United States: 2019*, provides the percentages of families in poverty. Suppose the 2017-2019[18] 3-year average percentage of families with female householder, no husband present, in poverty was 24.4. Suppose the percentages and bases for 2017, 2018, and 2019 were 26.2, 24.9, and 22.2 percent and 15,305,000, 15,052,000, and 14,838,000 respectively. Table 17 shows how to use the appropriate parameters and correlation coefficients from Tables 19 and 21 and Formulas (3), (16), and (17) to estimate the standard error and confidence interval.

Table 17. Illustration of Standard Errors of Data for Combined Years

	2017	2018	2019	2017-2019 Average
Percentage of families with female householder, no husband present, in poverty (*p*)	26.2	24.9	22.2	24.4
Base (*y*)	15,305,000	15,052,000	14,838,000	-
b-parameter (*b*)	1,518[A]	3,631[B]	5,529	-
Correlation (*r*)	-	-	-	0.35
Standard error	0.44	0.67	0.80	0.46
90-percent confidence interval	25.5 to 26.9	23.8 to 26.0	20.9 to 23.5	23.6 to 25.2

Source: U.S. Census Bureau, Current Population Survey, 2018-2020 Annual Social and Economic Supplement.

[A] This value comes from the Source and Accuracy Statement for the 2018 Annual Social and Economic Supplement, Appendix G, Table 19 in U.S. Census Bureau (2018). For additional information, see the "Year-to-Year Factors" section.

[B] This value comes from the Source and Accuracy Statement for the 2019 Annual Social and Economic Supplement, Appendix G, Table 19 in U.S. Census Bureau (2019d). For additional information, see the "Year-to-Year Factors" section.

[17] 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.

[18] The estimates for data year 2017 come from the CPS ASEC 2018 Bridge Files, and the estimates for data year 2018 come from the 2019 CPS ASEC Files.

The standard error of the 3-year average is calculated as

$$s_{\bar{x}} = \frac{1.37}{3} = 0.46$$

where

$$s_x = \sqrt{0.44^2 + 0.67^2 + 0.80^2 + (2 \times 0.35 \times 0.44 \times 0.67) + (2 \times 0.35 \times 0.67 \times 0.80)} = 1.37$$

The 90-percent confidence interval for the 3-year average percentage of families with a female householder, no husband present, in poverty is $24.4 \pm 1.645 \times 0.46$.

Standard Errors of Quarterly or Yearly Averages. For information on calculating standard errors for labor force data from the CPS which involve quarterly or yearly averages, please see Bureau of Labor Statistics (2006).

Year-to-Year Factors. In past years, the Census Bureau published a table of year factors for the CPS ASEC Supplement in the Source and Accuracy Statement. User demand for these factors has diminished with the introduction of replicate weights. Data users producing estimates from prior years should consult the Source and Accuracy Statements covering the years of their analysis to estimate standard errors.

Technical Assistance. If you require assistance or additional information, please contact the Demographic Statistical Methods Division via e-mail at dsmd.source.and.accuracy@census.gov.

Table 18. Parameters for Computation of Standard Errors for Labor Force Characteristics: March 2020

Characteristic	*a*	*b*
Total or White		
Civilian labor force, employed	-0.000013	2,481
Not in labor force	-0.000013	2,432
Unemployed	-0.000017	3,244
Civilian labor force, employed, not in labor force, and unemployed		
Men	-0.000031	2,947
Women	-0.000028	2,788
Both sexes, 16 to 19 years	-0.000261	3,244
Black		
Civilian labor force, employed, not in labor force, and unemployed	-0.000117	3,601
Men	-0.000249	3,465
Women	-0.000190	3,191
Both sexes, 16 to 19 years	-0.001425	3,601
Asian, American Indian and Alaska Native (AIAN), Native Hawaiian and Other Pacific Islander (NHOPI)		
Civilian labor force, employed, not in labor force, and unemployed	-0.000245	3,311
Men	-0.000537	3,397
Women	-0.000399	2,874
Both sexes, 16 to 19 years	-0.004078	3,311
Hispanic, may be of any race		
Civilian labor force, employed, not in labor force, and unemployed	-0.000087	3,316
Men	-0.000172	3,276
Women	-0.000158	3,001
Both sexes, 16 to 19 years	-0.000909	3,316

Source: U.S. Census Bureau, Internal Current Population Survey data files for the 2010 Design.

Notes: These parameters are to be applied to basic CPS monthly labor force estimates. The Total or White, Black, and Asian, AIAN, NHOPI parameters are to be used for both alone and in combination race group estimates. For same-sex households, multiply the a- and b-parameters by 1.3. For nonmetropolitan characteristics, multiply the a- and b-parameters by 1.5. If the characteristic of interest is total state population, not subtotaled by race or ethnicity, the a- and b-parameters are zero. 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 Black, Hispanic, and Asian, AIAN, NHOPI parameters. For the groups self-classified as having two or more races, use the Asian, AIAN, NHOPI parameters for all employment characteristics.

Table 19. Parameters for Computation of Standard Errors for People and Families: 2020 Annual Social and Economic Supplement

Characteristics	Total or White		Black		Asian, AIAN, & NHOPI[A]		Hispanic[B]	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*
PEOPLE								
Educational attainment	-0.000011	3,483	-0.000041	3,187	-0.000086	2,906	-0.000053	3,233
Employment	-0.000013	2,481	-0.000117	3,601	-0.000245	3,311	-0.000087	3,316
People by family income	-0.000019	6,000	-0.000075	5,781	-0.000142	4,820	-0.000089	5,403
Income characteristics								
Total	-0.000089	3,020	-0.000035	2,650	-0.000076	2,557	-0.000042	2,549
Male	-0.000081	2,736	-0.000074	2,685	-0.000155	2,538	-0.000097	2,952
Female	-0.000016	2,637	-0.000057	2,305	-0.000146	2,557	-0.000080	2,412
Age								
15 to 24	-0.000084	3,524	-0.000297	3,449	-0.000516	2,841	-0.000185	2,800
25 to 44	-0.000096	3,242	-0.000146	3,259	-0.000276	2,805	-0.000168	3,023
45 to 64	-0.000098	3,317	-0.000140	2,457	-0.000354	2,557	-0.000221	2,730
65 and over	-0.000061	3,270	-0.000249	2,193	-0.000741	2,686	-0.000487	2,324
Health insurance	-0.000009	3,022	-0.000034	2,598	-0.000095	3,223	-0.000060	3,633
Marital status, household and family								
Some household members	-0.000009	2,808	-0.000042	3,221	-0.000069	2,343	-0.000049	2,941
All household members	-0.000008	2,730	-0.000033	2,528	-0.000069	2,318	-0.000039	2,348
Mobility (movers)								
Educational attainment, labor force, Marital status, household, family, and income	-0.000013	4,135	-0.000054	4,181	-0.000104	3,505	-0.000063	3,841
US, county, state, region, or metropolitan statistical areas	-0.000018	5,986	-0.000066	5,104	-0.000137	4,629	-0.000095	5,734
Below poverty								
Total	-0.000113	3,838	-0.000108	3,667	-0.000092	3,099	-0.000106	3,572
Male	-0.000115	3,877	-0.000243	3,978	-0.000168	2,756	-0.000244	3,993
Female	-0.000107	3,603	-0.000206	3,589	-0.000182	3,183	-0.000210	3,659
Age								
Under 15	-0.000171	5,771	-0.000678	5,651	-0.000474	3,953	-0.000872	7,270
Under 18	-0.000112	3,781	-0.000420	4,341	-0.000310	3,202	-0.000435	4,495
15 and over	-0.000128	4,341	-0.000154	4,095	-0.000134	3,547	-0.000153	4,066
15 to 24	-0.000090	3,785	-0.000730	4,017	-0.000544	2,996	-0.000632	3,477
25 to 44	-0.000101	3,428	-0.000394	4,007	-0.000301	3,059	-0.000335	3,406
45 to 64	-0.000099	3,356	-0.000418	3,020	-0.000385	2,782	-0.000413	2,987
65 and over	-0.000063	3,395	-0.000714	2,588	-0.000909	3,294	-0.000701	2,538
Unemployment	-0.000017	3,244	-0.000117	3,601	-0.000245	3,311	-0.000087	3,316
FAMILIES, HOUSEHOLDS, OR UNRELATED INDIVIDUALS								
Income	-0.000030	3,938	-0.000184	3,930	-0.000261	3,420	-0.000134	3,866
Marital status, household and family, educational attainment, population by age/sex	-0.000009	3,238	-0.000066	2,550	-0.000285	3,754	-0.000074	3,758
Poverty	0.000103	5,529	0.000516	5,568	0.003231	3,933	0.000478	6,075

Source: U.S. Census Bureau, Current Population Survey, Internal data from the 2020 Annual Social and Economic Supplement.

[A] AIAN is American Indian and Alaska Native, and NHOPI is Native Hawaiian and Other Pacific Islander.
[B] Hispanics may be any race.
Notes: These parameters are to be applied to the 2020 Annual Social and Economic Supplement data. The Total or White, Black, and Asian, AIAN, NHOPI parameters are to be used for both alone and in combination race group estimates. For same-sex households, multiply the a- and b-parameters by 1.3. For nonmetropolitan characteristics, multiply the a- and b-parameters by 1.5. If the characteristic of interest is total state population, not subtotaled by race or ethnicity, the a- and b-parameters are zero. 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 Black, Asian, AIAN, NHOPI, and Hispanic parameters. For the group self-classified as having two or more races, use the Asian, AIAN, NHOPI parameters for all characteristics except employment, unemployment, and educational attainment, in which case use Black parameters. For a more detailed discussion on the use of parameters for race and ethnicity, please see the "Generalized Variance Parameters" section.

Table 20. Current Population Survey Year-to-Year Correlation Coefficients for Income and Health Insurance Characteristics: Data Years 1960 to 2019

Characteristics	1960-2000 (basic) or 2000 (expanded)-2019		1999 (basic)-2000 (expanded)	
	People	Families	People	Families
Total	**0.30**	**0.35**	**0.19**	**0.22**
White	0.30	0.35	0.20	0.23
Black	0.30	0.35	0.15	0.18
Other	0.30	0.35	0.15	0.17
Hispanic[A]	0.45	0.55	0.36	0.28

Source: U.S. Census Bureau, Current Population Survey, Internal data files.
[A] Hispanics may be any race.
Notes: Correlation coefficients are not available for income data before 1960. These correlation coefficients are for comparisons of consecutive years. For comparisons of nonconsecutive years, assume the correlation is zero. For households and unrelated individuals, use the correlation coefficient for families. For a more detailed discussion on the use of parameters for race and ethnicity, please see the "Generalized Variance Parameters" section.

Table 21. Current Population Survey Year-to-Year Correlation Coefficients for Poverty Characteristics: Data Years 1970 to 2019

Characteristics	1972-83, 1984-2000 (basic) or 2000 (expanded)-2019		1999 (basic)-2000 (expanded)		1983-1984		1971-1972		1970-1971	
	People	Families	People	Families	People	Families	People	Families	People	Families
Total	**0.45**	**0.35**	**0.29**	**0.22**	**0.39**	**0.30**	**0.15**	**0.14**	**0.31**	**0.28**
White	0.35	0.30	0.23	0.20	0.30	0.26	0.14	0.13	0.28	0.25
Black	0.45	0.35	0.23	0.18	0.39	0.30	0.17	0.16	0.35	0.32
Other	0.45	0.35	0.22	0.17	0.30	0.30	0.17	0.16	0.35	0.32
Hispanic[A]	0.65	0.55	0.52	0.40	0.56	0.47	0.17	0.16	0.35	0.32

Source: U.S. Census Bureau, Current Population Survey, Internal data files.

[A] Hispanics may be any race.

Notes: Correlation coefficients are not available for poverty data before 1970. These correlation coefficients are for comparisons of consecutive years. For comparisons of nonconsecutive years, assume the correlation is zero. For households and unrelated individuals, use the correlation coefficient for families. For a more detailed discussion on the use of parameters for race and ethnicity, please see the “Generalized Variance Parameters” section.

Table 22. Current Population Survey Correlation Coefficients Between Race and Subgroups: 2020 Annual Social and Economic Supplement

Race 1 (subgroup)	Race 2	*r*
White alone, not Hispanic	White alone	0.82
White alone, not Hispanic	White alone or in combination, not Hispanic	0.98
Black alone	Black alone or in combination	0.95
Asian alone	Asian alone or in combination	0.92

Source: U.S. Census Bureau, Current Population Survey, Internal data files.

Notes: For a more detailed discussion on the use of parameters for race and ethnicity, please see the “Generalized Variance Parameters” section.

Table 23. Factors and Populations for State Standard Errors and Parameters: 2020 Annual Social and Economic Supplement

State	Factor	Population	State	Factor	Population
Alabama	1.11	4,836,185	Montana	0.21	1,058,638
Alaska	0.18	703,401	Nebraska	0.52	1,910,003
Arizona	1.25	7,250,794	Nevada	0.77	3,077,543
Arkansas	0.73	2,968,859	New Hampshire	0.33	1,348,147
California	1.28	39,034,824	New Jersey	1.15	8,780,729
Colorado	1.22	5,707,954	New Mexico	0.51	2,062,715
Connecticut	0.86	3,516,977	New York	1.19	19,173,378
Delaware	0.22	964,590	North Carolina	1.18	10,353,123
District of Columbia	0.17	698,464	North Dakota	0.17	748,215
Florida	1.14	21,347,900	Ohio	1.10	11,524,840
Georgia	1.15	10,480,913	Oklahoma	1.06	3,886,392
Hawaii	0.32	1,356,765	Oregon	1.07	4,201,503
Idaho	0.41	1,790,518	Pennsylvania	1.11	12,603,961
Illinois	1.17	12,451,406	Rhode Island	0.28	1,044,437
Indiana	1.11	6,657,419	South Carolina	1.07	5,093,995
Iowa	0.77	3,116,100	South Dakota	0.22	870,562
Kansas	0.82	2,851,117	Tennessee	1.10	6,758,728
Kentucky	1.13	4,385,967	Texas	1.32	28,763,793
Louisiana	1.01	4,537,420	Utah	0.53	3,214,318
Maine	0.39	1,331,924	Vermont	0.18	617,810
Maryland	1.15	5,951,913	Virginia	1.19	8,345,522
Massachusetts	1.10	6,831,799	Washington	1.18	7,564,480
Michigan	1.11	9,883,888	West Virginia	0.48	1,755,736
Minnesota	1.13	5,604,353	Wisconsin	1.13	5,762,472
Mississippi	0.69	2,902,505	Wyoming	0.16	569,502
Missouri	1.13	6,035,560			

Source: U.S. Census Bureau, Current Population Survey, Internal data files for the 2010 Design; U.S. Census Bureau, Population Estimates, March 2020.

Notes: The state population counts in this table are for the 0+ population. For same-sex households, multiply the a- and b-parameters by 1.3. 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 Black, Asian, American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, and Hispanic.

Table 24. Factors and Populations for Regional Standard Errors and Parameters: 2020 Annual Social and Economic Supplement

Region	Factor	Population
Midwest	1.06	67,415,935
Northeast	1.07	55,249,162
South	1.13	124,032,005
West	1.12	77,592,955

Source: U.S. Census Bureau, Current Population Survey, Internal data files for the 2010 Design; U.S. Census Bureau, Population Estimates, March 2020.

Notes: The state population counts in this table are for the 0+ population. For same-sex households, multiply the a- and b-parameters by 1.3. 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 Black, Asian, American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, and Hispanic.

REFERENCES

Brooks, C.A. & Bailar, B.A. (1978). *Statistical Policy Working Paper 3 - An Error Profile: Employment as Measured by the Current Population Survey*. Subcommittee on Nonsampling Errors, Federal Committee on Statistical Methodology, U.S. Department of Commerce, Washington, DC. https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/spwp3.pdf

Bureau of Labor Statistics. (2006). *Household Data ("A" tables, monthly; "D" tables, quarterly)*. https://www.bls.gov/cps/eetech_methods.pdf

Bureau of Labor Statistics. (2014). *Redesign of the Sample for the Current Population Survey*. http://www.bls.gov/cps/sample_redesign_2014.pdf

Bureau of Labor Statistics. (2020). *The Employment Situation – March 2020*. https://www.bls.gov/news.release/archives/empsit_04032020.pdf

Rothbaum, J. & Bee, A. (2020). *Coronavirus Infects Surveys, Too: Nonresponse Bias During the Pandemic in the CPS ASEC*. https://www.census.gov/library/working-papers/2020/demo/SEHSD-WP2020-10.html

U.S. Census Bureau. (1978). *Money Income in 1976 of Families and Persons in the United States*. Current Population Reports, P60-114. Washington, DC: Government Printing Office. https://www2.census.gov/prod2/popscan/p60-114.pdf

U.S. Census Bureau. (1993). *Money Income of Households, Families, and Persons in the United States: 1992*. Current Population Reports, P60-184. Washington, DC: Government Printing Office. https://www2.census.gov/prod2/popscan/p60-184.pdf

U.S. Census Bureau. (2009). *Estimating ASEC Variances with Replicate Weights Part I: Instructions for Using the ASEC Public Use Replicate Weight File to Create ASEC Variance Estimates*. http://usa.ipums.org/usa/resources/repwt/Use_of_the_Public_Use_Replicate_Weight_File_final_PR.doc

U.S. Census Bureau. (2017). *Current Population Survey: 2017 Annual Social and Economic (ASEC) Supplement*. https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar17.pdf

U.S. Census Bureau. (2018). *Current Population Survey: 2018 Annual Social and Economic (ASEC) Supplement*. https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar18.pdf

U.S. Census Bureau. (2019a). 2017 CPS ASEC Research Files. https://www.census.gov/data/datasets/2017/demo/income-poverty/2017-cps-asec-research-file.html

U.S. Census Bureau. (2019b). *2018 CPS ASEC Bridge Files*. https://www.census.gov/data/datasets/2018/demo/income-poverty/cps-asec-bridge.html

U.S. Census Bureau. (2019c). *American Community Survey Accuracy of the Data (2018)*. https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/ACS_Accuracy_of_Data_2018.pdf

U.S. Census Bureau. (2019d). *Current Population Survey: 2019 Annual Social and Economic (ASEC) Supplement*. https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar19.pdf

U.S. Census Bureau. (2019e). *Current Population Survey: Design and Methodology*. Technical Paper 77. Washington, DC: Government Printing Office. https://www2.census.gov/programs-surveys/cps/methodology/CPS-Tech-Paper-77.pdf

U.S. Census Bureau. (2019f). *Health Insurance Coverage in the United States: 2018*. https://www.census.gov/content/dam/Census/library/publications/2019/demo/p60-267.pdf

U.S. Census Bureau. (2019g). *Survey Redesigns Make Comparisons to Years Before 2017 Difficult*. https://www.census.gov/library/stories/2019/09/us-median-household-income-not-significantly-different-from-2017.html

All online references accessed August 7, 2020.