Supporting Document A
Illustrative Examples of Standard

Version 1.2

Issued: 08 Jan 08

Census Bureau Standard
Dissemination of Census and Survey Data Products

Authored by:

Thomas F. Moore III
(Team Leader, DSMD)
Carol V. Caldwell (CSD)
Patrick J. Cantwell (SRD)
Eileen M. O'Brien (SRD)

USCENSUSBUREAU

Helping You Make Informed Decisions

Document Management & Control[1]

Version	Issue Date	Approval	Description
1.0	19 May 05	Associate Directors	Initial Release
1.1	09 Mar 06	Quality Mgr.	Inserted hyperlink for main standard.
1.2	08 Mar 08	Quality Mgr.	Added text to Requirement 4.0 to reflect revision to the standard.

[1] **The latest version of this document is maintained on the Census Bureau Intranet.**

Illustrative Examples of Standard

These examples are based on statements from actual reports. Out of context, they may appear to be incomplete, but their purpose is only to illustrate the requirement.

Requirement 1. All results, stated or implied, must be supported by the data.

Test results that are not statistically significant. When the test result for a published statement is not statistically significant, a statement denoting the lack of statistical significance is required.

1a.

Use the following blanket statement, for example: "All statements in this report have undergone statistical testing, and all comparisons are significant at the 90-percent confidence level unless otherwise noted." Doing so would then mean a statement in a report that "The median household income for the Northeast rose to $45,106" has been tested and the difference compared with the previous year is statistically significant.

If the statement continued to say, "The median household incomes for the remaining regions were $44,744 in the West, $44,646 in the Midwest, and $38,410 in the South, all statistically unchanged from the preceding year's income levels," then the changes for the other regions have been tested and each one's difference from the preceding year was not statistically significant.

Furthermore, as it stands, the statement also implies that all differences between regions in the current year have been tested and are statistically significant. If, for example, only the current year's value for the South is significantly different from the values for the other regions, the following footnote should be used: "The differences among the median household incomes for the Northeast, Midwest, and West regions

1b.

American Indians and Alaska Natives were less likely to have health insurance than other racial groups, based on a 3-year-average (1998-2000): 73.2 percent, compared with 80.5 percent of Blacks, 81.2 percent of Asians and Pacific Islanders, and 89.9 percent of non-Hispanic Whites. However, they were more likely to have insurance than were Hispanics (67.2 percent). Footnote: The difference in health insurance coverage rates between Blacks and Asians and Pacific Islanders was not statistically significant.

1c.

The difference between the unemployment rates of Black and Hispanic women was not statistically significant.

Alternative: Black women had an unemployment rate not statistically different from that of Hispanic women.

1d.

. . . Asian and Pacific Islander foreign born had a median length of residence of 13.6 years, and 45.7 percent reported they were naturalized citizens. Foreign-born Hispanics had a median length of residence of 13.2 years (not statistically different from the median for Asians and Pacific Islanders), but only 25.7 percent were naturalized citizens.

1e.

Median household income did not change statistically between 1999 and 2000, after experiencing 5 consecutive years of annual increases.

Alternative: After experiencing 5 consecutive years of annual increases, median household income is unchanged in 2000 (the difference from 1999 to 2000 is not statistically significant).

Differences. A difference compares a measurement A to a measurement B, where A and B represent different groups or the same group at different points in time. Comparing A to B is equivalent to comparing the difference A – B to 0, or the ratio A/B to 1. One approach is to determine whether the confidence interval for the difference A – B includes 0.

1f.

The 90-percent confidence interval around the difference is calculated as 0.5 to 2.1. Because this interval does not include zero, we can be reasonably confident that the percentage of women who are part-time workers is greater than the percentage of men who are part-time workers.

1g.

Foreign-born households experienced an increase in real median income between 1999 and 2000, but the income of native households did not change significantly.

Alternative: Foreign-born households experienced an increase in real median income between 1999 and 2000. The difference in income of native households during the same period, however, was not statistically significant.

1h.

> Between 1993 and 2000, Blacks experienced a 31.8 percent increase in their real per capita income, which rose to $15,197, up from $11,534 in 1993. This increase was larger than the increases for non-Hispanic Whites and Hispanics, but not statistically different from the increase experienced by Asians and Pacific Islanders.

Comparisons to 0. When stating that there is a change in a measurement, up or down, first test the change statistically as in the prior examples.

1i.

> Poverty rates fell for Blacks (from 23.6 percent to 22.1 percent) between 1999 and 2000.

Implicit statements about the data. Making statements about comparable measurements for two groups or two points in time allows the data user to calculate their difference. This is equivalent to making a statement about the difference between the estimates.

1j.

> Since 1993, family households have experienced a 17.4 increase in their median income and nonfamily households an increase of 14.5 percent. Footnote: There is no statistically significant difference between these estimates.

1k.

> Of the 9.0 million families in 2000 with a foreign-born householder, 75.6 percent were married-couple families, 16.7 percent were families with a female householder (no husband present), and 7.8 percent were families with a male householder (no wife present). Of the 63.0 million families with native householders, 77.0 percent were married-couple families, 17.8 percent were female householder families, and 5.2 percent were male householder families. *[Making this statement without further information implies there is a statistically significant difference between comparable estimates for foreign-born and native householders.]*

Multiple comparisons. When making simultaneous comparisons among a large number of groups or categories, take care to use appropriate statistical methods and to measure the resultant level of significance correctly. For example, performing many pair-wise statistical tests may lead to erroneous results and conclusions.

1l.

> Some comparisons required an adjustment to account for the effect of multiple comparisons in order to maintain an overall confidence level of 90 percent. The effect of using an adjustment factor is to make the individual comparison tests more conservative (i.e., less likely to detect a significant difference), while maintaining the overall confidence level.

Ranks. Make statements pertaining to the rank order of groups or categories (e.g., "The District of Columbia has the highest rate of college graduates") only after applying appropriate statistical procedures to ensure statistically significant differences exist between the groups or categories.

1m.

> Although the data may appear to suggest that Minnesota had the lowest uninsured rate, its estimated rate was not statistically different from the rates for Rhode Island, Wisconsin, or Iowa.

Presenting data in the absence of statistical testing. There may be situations where statistical testing is impractical or impossible because of the complexity of the inference. Observations about the data may be made provided they don't state or imply any statistically unsupported inferences.

1n.

> About one in eight single-family homes in Cambridge, Mass., were valued at $1 million or more in 2000, the highest proportion in the nation among large cities, according to the Commerce Department's Census Bureau.
>
> San Francisco (7.0 percent), Pasadena, Calif., (4.7 percent) and Los Angeles (3.8 percent) rounded out the top four in percentage of $1 million homes among cities of 100,000 population or more. The next six cities were in the neighborhood of 3 percent (see table).
>
> Note [in the table]: The estimates may not be significantly different from one another or from estimates for other geographic areas not listed in these tables.

1o.

> Of the foreign-born population in 2000, Asian and Latin American countries accounted for 9 of the top 10 leading countries of birth. However, the Current Population Survey's sample size is not large enough to rank most countries accurately. See the 90-percent confidence intervals shown in Figure _.

Requirement 2. The level of detail for a published table must be appropriate for the sampling and nonsampling errors associated with the estimates.

2a.

> Despite its recent expansion to 78,000 households nationwide, the Annual Social and Economic Supplement to the Current Population Survey is not large enough to produce reliable annual estimates for American Indians and Alaska Natives. However, Table 2 displays 3-year averages of the number of American Indians and Alaska Natives in poverty and their 3-year-average poverty rate and, for comparison, provides 3-year-average poverty statistics for the other groups.

2b.

S Estimate suppressed because tests for identifiable and stable seasonality do not meet reliability standards.

2c.

The symbol "B" is used to indicate a derived figure that was suppressed because the base for the derived figure is less than 75,000.

Requirement 3. All data products or news releases should include text stating that the data are subject to error arising from a variety of sources.

Sampling and nonsampling errors. Possible statements to describe the sources of errors are the following:

3a.

See [website address] for information on sampling and nonsampling errors in the Commodity Flow Survey.

3b.

Sampling errors occur because observations are made on a sample, not on the entire population, and different samples could have led to different estimates. Standard errors and coefficients of variation, as calculated for this report, are measures of sampling variation. The margin of error, as used on page _, gives a range around the estimate that is a 90% confidence interval. If, for example, the percent change estimate is +1.2% and the standard error is 0.9%, then the margin of error is ± 1.65 x 0.9% or ± 1.5%, and the 90% confidence interval is -0.3% to +2.7%. If the interval contains 0, as it does here, then we do not have sufficient evidence to conclude that the estimated change was statistically different from 0.

3c.

Nonsampling errors may arise due to [any of the following]: lack of coverage, duplication in the frame, unit or item nonresponse, imputation, response variability, response bias, interviewer effects, problems with the mode of collection, or processing error.

3d.

A sample survey has two possible types of error: sampling and nonsampling. The accuracy of an estimate depends on both types of error, but the full extent of the nonsampling error is generally unknown. The standard errors for CPS estimates primarily indicate the magnitude of sampling error. They also partially measure the effect of some nonsampling errors in responses and enumeration, but do not measure systematic biases in the data. (The bias is the difference between the average of the sample estimate over all possible samples and the true value.)

Nonsampling Error. Other sources of error in the survey estimates are collectively called nonsampling error. Sources of nonsampling error include the following:

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

The Census Bureau employs quality control procedures throughout the production process, including the overall design of surveys, the wording of questions, the review of the work of interviewers and coders, and the statistical review of reports to minimize these errors.

3e.

Census questionnaires were sent to all known wholesale trade and retail trade employers and therefore, are not subject to sampling variability. However, the data are subject to nonsampling errors. Nonsampling errors can be attributed to many sources: inability to identify all cases in the actual universe; inability or unwillingness on the part of respondents to provide correct information; definition and classification difficulties; response errors and bias; errors in collection or processing; misinterpretation of questions; and other errors of recording, keying, and estimation for missing or misreported data.

The accuracy of these tabulated data is influenced by the joint effects of the various nonsampling errors. Explicit measures of the effects of these nonsampling errors are not available. However, it is believed that most of the important operational and response errors were detected and corrected through systematic clerical edits, automated data edits, and an analyst review.

3f.

The Census Bureau employs quality control procedures throughout the production process, including the overall design of surveys, the wording of questions, the review of the work of interviewers and coders, and the statistical review of reports to minimize these errors.

3g.

Data in this report are based on a sample and, therefore, are subject to sampling and nonsampling error. A discussion of the reliability of the estimates and general survey methodology appears in BW/01-A Current Business Reports.

3h.

The estimates in the income and poverty reports are based on the 2001, 2002 and 2003 Annual Social and Economic Supplement to the Current Population Survey. As in all surveys, the data are subject to sampling variability and other sources of error.

3i.

The tabulations in the report are based on responses from the sample of households that received the census long form, about 1-in-6 nationally. They are subject to sampling and nonsampling error.

Requirement 4. Make estimates of the magnitude of sampling error available for all survey estimates and comparisons. Include 90-percent confidence intervals with the point estimates for key survey estimates and comparisons.

4a.

Total retail sales for the second quarter were estimated at $858.8 billion, an increase of 4.9% (± 0.3%) from the same period a year ago.

4b.

The nation's poverty rate rose from 11.3 percent in 2000 to 11.7 percent in 2001. Footnote: The margin of error for the change in the nation's poverty rate from 2000 to 2001 is approximately ± 0.2%, giving a range of 0.2% to 0.6%.

4c.

Reliability of Key Estimates

Estimate	2000	2001	Estimated change	Lower bound	Upper bound
Poverty rate	11.3%	11.7%	0.4%	0.2%	0.6%
Median household income	$43,162	$42,228	-$934	-$1182	-$686

4d.

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

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

Here x is the size of the estimate and a and b are the parameters in Table _ associated with the particular type of characteristic. When calculating standard errors for numbers from cross-tabulations involving different characteristics, use the factor or set of parameters for the characteristic which will give the largest standard error.

Authors of *press releases* are encouraged to include confidence intervals or margins of error for key estimates, as shown in example 4e.

4e.

Real median household income in the United States climbed between 2005 and 2006, reaching $48,200, according to a report released today by the U.S. Census Bureau. This is the second consecutive year that income has risen. The margin of error is plus or minus $341.

Meanwhile, the nation's official poverty rate declined for the first time this decade, from 12.6 percent in 2005 to 12.3 percent in 2006. There were 36.5 million people in poverty in 2006, not statistically different from 2005. The number of people without health insurance coverage rose from 44.8 million (15.3 percent) in 2005 to 47 million (15.8 percent) in 2006. The margin of error on the poverty rate is plus or minus 0.2 percent and on the number without health insurance, plus or minus one-half million people.

Absolute vs. relative standard errors. Standard errors may be presented in either absolute or relative terms. Relative error refers to the coefficient of variation (estimated by the ratio of the standard error of the estimate to the estimate itself) or its square, known as the relative variance. When the estimates are percentages or proportions, providing an estimate of the magnitude of the absolute error is generally preferred over that of the relative error. It should be clear to the user which type of error is being measured.

4f.

For monthly totals, the coefficient of variation (CV) is given. The resulting 90-percent confidence interval is the estimated value ± 1.65 × CV × (the estimated

Requirement 5. Present indicators of potential nonsampling error.

5a.

Nonresponse. The effect of nonresponse cannot be measured directly, but one indication of its potential effect is the nonresponse rate. For the March 2001 basic Current Population Survey (CPS), the nonresponse rate was 8.03%. The nonresponse rate for the supplement was an additional 8.50%, for an effective combined supplement nonresponse rate of 15.85%.

5b.

Coverage. The concept of coverage in the survey sampling process is the extent to which the total population that could be selected for sample "covers" the survey's target population. CPS undercoverage results from missed housing units and missed people within sample households. Overall CPS undercoverage is estimated to be about 8 percent. CPS undercoverage varies with age, sex, race, and Hispanic origin. Generally, undercoverage is larger for males than for females, and larger for Blacks and other races combined than for Whites. As described previously, ratio estimation to independent age-sex-race-Hispanic population controls partially corrects for the bias due to undercoverage. However, biases exist in the estimates to the extent that missed people have different characteristics from those of interviewed people in the same age-sex-race-Hispanic group.

A common measure of survey coverage is the coverage ratio, the estimated population before ratio estimation divided by the independent population control. Table 1 shows CPS coverage ratios for age-sex-race groups for a typical month. The CPS coverage ratios can exhibit some variability from month to month. Other Census Bureau household surveys experience similar coverage.

Table 1. CPS Coverage Ratios

	Non-Black		Black		All People		
Age	M	F	M	F	M	F	Total
0-14	0.929	0.964	0.850	0.838	0.916	0.943	0.929
15	0.933	0.895	0.763	0.824	0.905	0.883	0.895
16-19	0.881	0.891	0.711	0.802	0.855	0.877	0.866
20-29	0.847	0.897	0.660	0.811	0.823	0.884	0.854
30-39	0.904	0.931	0.680	0.845	0.877	0.920	0.899
40-49	0.928	0.966	0.816	0.911	0.917	0.959	0.938
50-59	0.953	0.974	0.896	0.927	0.948	0.969	0.959
60-64	0.961	0.941	0.954	0.953	0.960	0.942	0.950
65-69	0.919	0.972	0.982	0.984	0.924	0.973	0.951
70+	0.993	1.004	0.996	0.979	0.993	1.002	0.998
15+	0.914	0.945	0.767	0.874	0.898	0.927	0.918
0+	0.918	0.949	0.793	0.864	0.902	0.931	0.921

5c.

A Nonsampling Error Warning. Because the full extent of the nonsampling error is unknown, one should be particularly careful when interpreting results in which the relevant difference or statistic is near the margin of statistical significance. Even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test. Caution should also be used when interpreting results based on a relatively small number of cases. Summary measures probably do not reveal useful information when based on fewer than 75,000 people.

For additional information on nonsampling error, including its potential effect on CPS data when known, refer to:

- Statistical Policy Working Paper 3, *An Error Profile: Employment as Measured by the Current Population Survey*, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce, 1978
- Technical Paper 63, *Current Population Survey: Design and Methodology*, U.S. Census Bureau, U.S. Department of Commerce, 2000.

5d.

The basic CPS nonresponse rates are for households; nonresponse rates for the supplement are for people. In order for an individual to be *eligible* for the supplement, an individual's CPS questionnaire must have been *completed*. In the nonresponse rates for 'Supplement - Total,' proxy interviews are treated as interviews; in the nonresponse rates for 'Supplement - Self Only,' proxy responses are treated as noninterviews.

5e.

Nonsampling errors can occur because of nonresponse, insufficient coverage of the universe of wholesale businesses, mistakes in the recording and coding of data, and response errors. Imputed data account for approximately 28 percent of the sales estimates and 31 percent of the inventories estimates. Additionally, estimates of sales and inventories prior to March 2001 have been restated from estimates based on the Standard Industrial Classification (SIC). The restatement methodology may have introduced additional nonsampling error. The calculated median standard errors of year-to-year change estimates may also be influenced by the methodology used to impute historical data for units in the sample based on the North American Industry Classification System (NAICS). The effect of this historic imputation methodology on published estimates will decrease as more data from the NAICS sample is accumulated. Precautionary steps are taken to minimize nonsampling errors, but their magnitude is not directly measured.

5f.

Item Nonresponse. Because inserting (imputing) values for missing data can introduce error into the estimates, the item nonresponse rate can be used to help characterize the nonsampling error attributed to item nonresponse. The item nonresponse rates from the 1999 National Survey of College Graduates (NSCG) were calculated using the simple imputation rate formula. The simple imputation rate calculates the percentage of "useable" data that was imputed for each questionnaire item. [The formula may be shown here.]

The table below provides the unweighted and weighted simple imputation rates by questionnaire section and for the overall 1999 NSCG.

1999 National Survey of College Graduates
Item Nonresponse Rates

1999 NSCG Questionnaire Section	Item Nonresponse Rate	
	Unweighted	Weighted
Section A: Employment Status	1.84%	1.92%
Section B: Past Employment	0.06%	0.06%
Section C: Other Work-Related Information	2.02%	2.11%
Section D: Background Information	0.82%	0.85%
Overall 1999 NSCG Totals	1.48%	1.55%