Multi-Year Estimate Study Quality Measures Definitions

Introduction

The Multi-Year Estimate Study consists of profile summaries that display the estimate and percent distributions of selected characteristics measured by the ACS during the different time periods. Each type of estimate is included in these profiles accompanied by its margin of error.

This document describes the definitions and computation of four quality measures:

1. Sample Size
2. Response Rates and Reasons for Noninterviews
3. Coverage Rates
4. Item Allocation Rates

The quality measures provide additional information about the quality of the data. Items related to all four measures are discussed in Common Methodology. Descriptions of each quality measure are then discussed individually. The definitions in these sections are written as Question and Answers. To learn more about the quality measures data files and the variables on the files, see the document titled "ACS Multi-Year Estimates Study Quality Measures File Layouts" available on the FTP site with filenames "myeqmfiles.doc" or "myeqmfiles.pdf".

Common Methodology

1. What time periods are covered for the quality measures?

The sample size measures are produced for one-, three-, and five-year time periods.

There are six one-year (12 months) time periods: 2000, 2001, 2002, 2003, 2004 and 2005.

There are five three-year (36 months) time periods: 1999-2001, 2000-2002, 2001-2003, 2002-2004, and 2003-2005.

There are three five-year (60 months) time periods: 1999-2003, 2000-2004, and 2001-2005.

2. Are the quality measures calculated for different types of geographies for the one-, three-, and five-year time periods?

The quality measures were produced for all geographic areas that are tabulated down to the Census tract level. No quality measures are produced for Census block groups. Coverage rates are only produced at the county level. The following gives more specific information about the different period estimates.

The quality measures for the one-year time periods are calculated for counties, minor civil divisions, place parts, PUMAs, secondary school districts, and unified school districts that meet the 65,000 threshold. The elementary school districts and American Indian areas did not meet

the required population threshold for any of the one-year time periods, otherwise these would be included in the measures.

The quality measures for the three-year time periods are calculated for counties, minor civil divisions, place parts, PUMAs, elementary school districts, secondary school districts, and unified school districts that meet the 20,000 threshold. The American Indian areas did not meet the required population threshold for any of the three-year time periods, otherwise these would be included in the measures.

The quality measures for the five-year time periods are calculated for counties, minor civil divisions, place parts, American Indian areas, PUMAs, elementary school districts, secondary school districts, unified school districts, census tracts, and five-digit zip codes.

3. What geographic boundaries are used for the three- and five-year time periods?

The geographic boundaries, such as state, county, and tract are based on the final year of the time period. For example, the 2000-2004 time period would use the 2004 geographic boundaries.

Sample Size -- Definitions

1. Do the sample size measures use tabulation geography or time of sampling geography?

The tabulation geography is used for all sample size measures.

2. What is sampling error?

ACS estimates are based on data from a sample rather than from all units in the population. For this reason, ACS estimates have a degree of uncertainty associated with them, called sampling error. In general, the larger the sample, the smaller the level of sampling error.

3. Why is it important to measure sampling error?

The estimates produced by the ACS are not exact because they are based on a sample. The sampling error measures the degree of uncertainty associated with the estimate. If the degree of uncertainty is too large, users should be cautious in how the estimates are used.

4. How does the ACS measure sampling error?

The ACS calculates standard errors for each estimate produced and publishes the 90 percent confidence level margin of error (the Census Bureau standard). You can be 90 percent confident that the interval that has the margin of error width on each side of the estimate includes the true value. See Accuracy of the Data for more details on how the margin of error and confidence intervals are calculated and interpreted. Note that the margin of error does not include the effect of nonsampling error (which is discussed in a later section).

5. What other numbers provide important information on the reliability of ACS estimates?

Initial Addresses Selected - The initial addresses selected is the number of addresses that were selected for the ACS sample for one-, three-, or five-year time periods. This includes addresses later determined to be commercial or nonexistent, as well as housing units that are not interviewed due to subsampling for personal visit follow-up, refusals or other reasons.

Final Interviews - The final number of interviews across all three modes of data collection for the ACS for one-, three-, or five-year time periods. This number includes occupied and vacant housing units that were interviewed by mail, telephone, or personal visit methods during each year in the time period. It excludes addresses determined to be nonexistent or commercial, and addresses not selected in the sample for personal visit follow-up, and addresses that are not interviewed due to refusals or other reasons.

6. How can the number of interviews exceed the number of initially selected addresses for a given geography and period?

The number of initially selected addresses for a given geography and period is based on the sample that was selected during that time period. The number of interviews is based on the interviews collected during that same time period. While most of the interviews for a given multi-year period were sampled in that same period, because of the data collection methods used by the ACS, the housing unit could have been selected in sample in a year preceding the period but was interviewed in the given period. Likewise, some addresses that are selected for sample in a given period may be interviewed in the year following the period. For most geographies, this detail goes unnoticed but for very small geographies, the data user may find instances where there is a smaller number of sample addresses for the period than interviewed housing units.

7. Are these numbers used to calculate the survey response rate?

These numbers by themselves are not used to measure the response to the request to participate in the ACS. See the section on Response Rates and Reasons for Noninterviews -- Definitions for more information on the calculation of the survey response rate.

8. Was each county sampled at the same rate for each year in the study?

The sampling rate by county varied from year to year during the years covered by this study. In general, the sampling rates were higher in most counties during 1999-2001 than in the other years with a few exceptions. For a complete overview of the sampling rate determination by year, please see the Accuracy of the Data.

9. How does the ACS calculate these two measures of sample size?

Initial Addresses Selected for geography x and period z =

Sum of the monthly address samples selected from the Master Address File for geography x and period z.

Final Interviews for geography x and year z =

Sum of all interviews successfully conducted by mail, telephone, or personal visit for geography x and period z.

Response Rates and Reasons for Noninterviews -- Definitions

1. What is Unit Nonresponse?

Unit nonresponse is the failure to obtain the minimum required information from a housing unit in the sample. Unit nonresponse occurs when respondents are unable or unwilling to participate, interviewers are unable to locate addresses or respondents, or when other barriers exist to completing the interview.

2. How does the ACS adjust for unit nonresponse?

The ACS uses noninterview adjustment methods to give a higher weight to interviewed units. Weight adjustments are made at the tract level within each sample month. See Accuracy of the Data for more information on how the noninterview adjustment is calculated.

3. Why is it important to measure unit nonresponse?

We measure it because it has a direct effect on the quality of the data. If the rate of unit nonresponse is high, it increases the chance that the final survey estimates may reflect bias. Estimates may reflect bias if the characteristics of nonresponding units differ from the characteristics of responding units. Unit nonresponse can contribute to nonsampling error.

4. How does the ACS measure unit nonresponse?

The Census Bureau calculates survey response rates to measure unit nonresponse in the ACS. The survey response rate is the ratio of the estimate of housing units interviewed after data collection is complete to the estimate of all units that should have been interviewed. For the ACS, this means all interviews after mail, telephone and personal visit follow-up. Interviews include complete and partial interviews with enough information to be processed.

To accurately measure unit nonresponse the ACS must estimate the universe of cases eligible to be interviewed and the survey noninterviews; that is, all eligible units in personal visit follow-up are given the appropriate weight as are all the noninterviews.

5. What are the primary reasons for unit nonresponse in the ACS?

The Census Bureau classifies all final noninterviews by one of the following seven **Reasons for Noninterviews** to understand why unit nonresponse occurred:

- **Refusal:** Even though the ACS is a mandatory survey and households whose addresses are selected for the survey are required to answer the survey questions, a few households may feel reluctant to cooperate and refuse to participate.

- **Unable to Locate:** If the interviewer cannot find the sample address after using all possible sources, they consider it "unable to locate".
- **No One Home:** Interviewers assign this code if they could not find anyone at the housing unit during the entire month's interview period.
- **Temporarily Absent:** The interviewer confirms that all household members are away during the entire month's interview period, for example on vacation, on a business trip, or caring for sick relatives.
- **Language Problem:** The interviewer could not conduct an interview because of language barriers, was not able to get an interpreter who could translate, and the supervisor or regional office could not help complete this case.
- **Insufficient Data:** To be considered an interviewed unit in ACS, a household's response must include a minimum amount of answers for basic survey items for the occupants, such as age and sex. Occupied housing units not meeting this minimum standard are treated as noninterviews in the estimation process. Responses for vacant housing units are not subject to a minimum data requirement.
- **Other:** Unique situations when the reason for noninterview does not fit into one of the classifications described above. Possible reasons include "death in the family", "household quarantined", or "roads impassable".

6. Why does the Census Bureau weight the survey response rate?

Weighting is used because not all housing units have the same probability of selection. The ACS sample design includes differential sampling and subsampling rates. (Refer to the Accuracy of the Data for more information about the use of differential sampling and subsampling in the ACS.) Weighting accounts for the different probabilities of selection as a result of sampling and subsampling.

7. Why is there a spike in the Other noninterview rates in 2002 and 2004?

In 2002 and 2004, due to budget constraints, we were unable to interview or follow-up certain addresses that had been sampled. An entire sample month of selected addresses was not mailed forms or followed up in 2002, and in 2004 the nonrespondents from one month's mail phase were not followed up by telephone or personal interview. These cases were treated as Other noninterviews for purposes of calculating these quality measures.

8. How are survey response rates calculated?

Response rate for geography x and period z =

$$\left(\frac{\text{Initially weighted estimate of interviews for geography x and period z}}{\text{Initially weighted estimate of cases eligible to be interviewed for geography x and period z}} \right) \times 100$$

Noninterview rate for reason A for geography x and period z =

$$\left(\frac{\text{Initially weighted estimate of cases not interviewed for reason A for geography x and period z}}{\text{Initially weighted estimate of cases eligible to be interviewed for geography x and period z}}\right)\times 100$$

Coverage Rates -- Definitions

1. What is coverage error?

There are two kinds of coverage error: under-coverage and over-coverage.

Under-coverage exists when housing units or people do not have a chance of being selected in the sample.

Over-coverage exists when housing units or people have more than one chance of selection in the sample, or are included in the sample when they should not have been.

2. How does the ACS reduce coverage error?

The final ACS population estimates are adjusted for coverage error by controlling specific survey estimates to independent population controls by sex, age, race, and Hispanic origin. The final ACS estimates of housing units are controlled to independent estimates of total housing.

Refer to Accuracy of the Data to learn more about this weighting procedure.

3. Why is it important to measure coverage error?

If the characteristics of under-covered or over-covered housing units or individuals differ from those that are selected, the ACS may not provide an accurate picture of the population. Under- and over-coverage are also sources of nonsampling error.

4. How does the ACS measure coverage error?

The Census Bureau calculates coverage rates to measure coverage error in the ACS. The coverage rate is the ratio of the ACS population estimate of an area or group to the independent estimate for that area or group, times 100.

Coverage rates are calculated for the total population and by sex at the county level. These rates are weighted to reflect the probability of selection into the sample and subsampling for personal visit follow-up. As the coverage rate drops below 100 percent, the weights of the people in the survey need greater adjustment in the final weighting procedure to reach the independent estimate. If the rate is greater than 100 percent, the ACS population estimates are downweighted to match the independent estimates.

5. What is the independent population estimate?

The Census Bureau uses independent data on births, deaths, immigration, etc. to produce official estimates of the population each year. The base for these independent estimates is the decennial census counts. The ACS, like all other Census Bureau household surveys, controls some of its most basic estimates to these official estimates to correct for potential over- or under-coverage.

6. How are Coverage Rates calculated?

Coverage rate for county x and period z =

$$\left(\frac{\text{uncontrolled ACS estimate of total persons for county x and period z}}{\text{official estimate of total persons for county x and period z}} \right) \times 100$$

Coverage rate for group A for county x and period z =

$$\left(\frac{\text{uncontrolled ACS estimate of total persons for group A for county x and period z}}{\text{official estimate of total persons for group A for county x and period z}} \right) \times 100$$

Item Allocation Rates -- Definitions

1. What is item nonresponse?

Missing data for a particular question or item is called item nonresponse. It occurs when a respondent fails to provide an answer to a required item. The ACS also considers invalid answers as item nonresponse.

2. How does the ACS correct for item nonresponse?

The Census Bureau uses imputation methods that either use rules to determine acceptable answers or use answers from similar housing units or people who provided the item information. The first of these two methods is known as "assignment," while the second is referred to as "allocation."

Assignment involves logical imputation where a response to one question implies the value for a missing response to another question. For example, first name can often be used to assign a value to sex.

Allocation, on the other hand, involves using statistical procedures, such as within-household or nearest neighbor matrices populated by donors, to impute for missing values.

3. Why is it important to measure item nonresponse?

So data users can judge the completeness of the data in which the survey estimates are based. Final estimates can be adversely impacted when item nonresponse is high and bias can be introduced if the characteristics of the nonrespondents differ from those reported by respondents. Item nonresponse and unit nonresponse both contribute to potential bias in the estimates, and are sources of nonsampling error.

4. How does the ACS measure item nonresponse?

Item nonresponse is measured through the calculation of allocation rates which are published with the survey estimates.

5. How are item allocation rates calculated?

Allocation rate for item A for geography x and period z =

$$\left(\frac{\text{total number of weighted responses allocated for item A for geography x and period z}}{\text{total number of weighted responses for item A for geography x and period z}} \right) \times 100$$