American Community Survey

Data Suppression

September 27, 2016

Table of Contents

1   Introduction

Data users often question why certain American Community Survey (ACS) estimates are not available. The missing estimates can be caused by data suppression. Data suppression refers to the various methods or restrictions that are applied to ACS estimates to limit the disclosure of information about individual respondents and to reduce the number of estimates with unacceptable levels of statistical reliability.

Section 2 of this document explains the factors that contribute to data suppression.

2   Data Suppression Factors

2.1   Population Thresholds

The ACS population thresholds apply to 1-year and Supplemental Estimates only. If a geographic area has a population at or exceeding the population threshold, the estimates for that area will be published. Table 1 below shows the population thresholds for ACS estimates. As shown by this table, areas or groups of 65,000 or more are eligible for 1-, 5-year, and Supplemental Estimates. Areas or groups of 20,000 or more are eligible for 5-year and Supplemental Estimates. Areas or groups of 20,000 or fewer are eligible for 5-year estimates only.

Estimate Type	Population Threshold
1-Year Estimates	65,000
Supplemental Estimates	20,000
5-Year Estimates	Not applicable

Table 1: ACS Population Thresholds by Estimate Type

Population estimates from the Census Bureau Population Estimates Program (PEP) are used to determine the availability of data for geographic areas such as state, county, sub-county and place. Otherwise, ACS estimates are used.

There are two exceptions to the above rule:

- One-year ACS estimates are published for all Metropolitan Statistical Areas (MSAs) with populations of at least 50,000 based on ACS estimates. This exception is designed to acknowledge the importance of the MSA standard as a basis for collecting, tabulating, and publishing Federal statistics.

- The 1-year ACS estimates will be published if the area's total population falls within 5% of the 65,000 and the area was published for previous year. This exception to the population threshold rules shown in Table 1 is popularly known as the "grandfather rule". It is designed to provide a slight relaxation to the standard ACS data publication thresholds in the interest of providing as much ACS data as possible for local areas. It also allows for year-to-year changes in community size that could have resulted from unusual events such as natural disasters that caused local populations to decline temporarily.

There are also special rules for the Selected Population Profiles.

- For the 1- and 3-year ACS, there must be at least 500,000 people in the geographic area.
- There must be at least 65,000 people (for 1-year) and 7,000 people (for 5-year) in the specific population subgroup. There are no Selected Population Profiles in the Supplemental Estimates.

2.2   **Geographic Restrictions for Tables**

The subject matter areas consider data reliability, rules from the Census Bureau Disclosure Review Board (DRB), and the nature of the topics covered by the table making a determination of the appropriate geographic areas that each table is published for. Table 2 shows the major reasons for geographic restrictions of tables. For a full list of table geographic restrictions, see Appendix B: List of Geographic Restrictions by Table

Restriction Reason	Number of Tables	Examples	Published Geographic Areas
Data Reliability	30	Table B02003 (Detailed Race) contains estimates for 60 plus detailed race groups.	Nation
		Table B12501 (Marriages In The Last Year By Sex By Marital Status For The Population 15 Years And Over) contains detailed marital history for only the current year.	Nation and States
Disclosure Review Board Rules	445	Table B24122 (Detailed Occupation By Median Earnings In The Past 12 Months For The Full-Time, Year-Round Civilian Employed Male Population 16 Years And Over) contains 526 detailed occupation categories.	No Block Groups
		Table B05003 (Sex By Age By Nativity And Citizenship Status) contains foreign-born estimates broken down by citizenship status.	No Block Groups

| Nature of the Topic | 218 | Table B07101 (Movers Between Regions In The United States) is designed for high level migration estimates. | Nation |
| | | Table B08406 (Sex Of Workers By Means Of Transportation To Work For Workplace Geography) measures commuting patterns. | Geographic Areas below State Level |

Table 2: Geographic Restrictions for Tables

2.3 **Data Quality Filtering**

Data quality filtering rules are applied to ACS 1-year and Supplemental Estimates, but are not applied to the 5-year estimates. Every detailed table consists of a series of estimates and each estimate is subject to sampling variability that can be summarized by its standard error. If more than half of the estimates in the table are not statistically different from 0 (at a 90 percent confidence level), then the table fails the test and is restricted from publication. Below are the detailed filtering rules.

- For detailed tables, a table is filtered out if the median CV is greater than 0.61. A cell with an estimate of 0 is considered to have a CV = 1. Only lines at the lowest level of detail of a table are included in the calculation of the median CV.

- Ratio tables derived from other detailed tables are filtered out if the numerator or denominator estimate from the detailed tables is filtered out.

- Cells in data profiles, subject tables and ranking/GCT tables are filtered out if the data from detailed tables that is used to source the cell is filtered out. Table with means, unlike detailed tables, can have some lines filtered out and some lines not filtered out.

- Subject tables featuring specific population groups and iterated Selected Population Profiles (SPPs) that are generated directly from microdata are filtered in a similar matter to detailed tables, except that filtering is applied to the sub-population groups as opposed to the whole tables. However, if one half or more of the lines in a SPP for a geography are filtered out, the whole table is filtered out for that geography.

2.4 **Collapsed Tables**

ACS produces and releases two types of detailed tables: base tables and collapsed tables. Base tables provide the most detailed estimates. Collapsed tables represent a less detailed version of the base tables achieved by collapsing categories and combining estimates into fewer categories. The collapsed tables were introduced in an attempt to release more data to the public because many of the corresponding base tables were failing data quality filtering. The collapsed tables

are smaller tables with fewer details than its base table counterparts. As a result, collapsed tables generally have a better chance of passing data quality filtering and being released.

The first letter of "C" in the Table ID identifies collapsed tables, while the first letter of "B" identifies its corresponding base tables. Table 3 below shows the examples of the "B" and "C" tables. Please note that due to the size of the product, only the "B" or "C" table is released for a specific table in the 5-year datasets.

Base Table	Collapsed Table	Table Description
B01001A	C01001A	Sex by Age (White Alone)
B01001D	C01001D	Sex by Age (Asian Alone)
B02006	C02006	Asian alone by selected groups
B21005	C21005	Age by veteran status by employment status for the civilian population 18 to 64 years

Table 3: Examples of "B" and "C" Tables

2.5   Disclosure Review Board (DRB) Rules

Beyond the population thresholds applied to ACS estimates and data quality filtering applied to ACS tables, the Disclosure Review Board (DRB) establishes additional rules that specify what ACS data are released to ensure that the confidentiality of the respondents' data are protected. Proposed new and modified ACS tables will be submitted to DRB for review and approval each calendar year before production begins.

See Appendix A: DRB Memo for details of the rules on 2012 ACS estimates. Below outlines the summary of the rules.

2.5.1   Minimum Number of Cases

- Non-residence geography (place of work, place of birth, residence one year ago): 50 or more cases in the geographic area

- Selected Population Profiles: 50 or more cases in the geographic area

- Unweighted number of housing units/population: 3 or more occupied housing unit cases in the geographic area

- Means and Aggregates:
    - DRB rules:
        - 3 or more cases are required to publish a cell in a table
        - Complementary suppression is required to prevent users from calculating suppressed data

▪ ACS practice:
  • Suppress the whole table if any cell does not meet the standard

2.5.2   Table Cells and Table Topics Restrictions

Table with more than 100 independent cells or tables with sensitive topics that could lead to disclosure risks are not allowed at block groups. Some of the examples are tables for Place of Birth for the Foreign-Born Population; Sex by Occupation for the Civilian Employed Population 16 Years and Older; Sex by Age, only when iterated by race/Hispanic origin; Characteristics of non-citizens; and specific type of disability by race.

2.6   Table Suppression for Specific Geographies

During the data and data products review period, subject matter analysts may identify microdata errors, which occurred during data collection or early data processing. These errors may have resulted in unreasonable estimates appearing in tabulated tables for associated geographic areas.

When ACS management determines it is not feasible to correct data errors based on an evaluation of the data release schedule against risks and resources, the method used to avoid releasing unreliable estimates is to suppress the affected tables for specific geographic areas and thereby prevent their release to the public. The suppression that is recommended must be approved by the ACS Change Control Board. Some examples of tables suppressed for the 2011 ACS data release are shown in Table 4.

Table ID	Table Description	Area Name	Data Year
B25048	Plumbing facilities for occupied housing units	Arecibo Municipio, Puerto Rico	2011
B25043	Tenure by telephone service available by age of householder	Dodge County School District, Georgia	2011
C25016	Tenure by plumbing facilities by occupants per room	Aguadilla-Isabela-San Sebastián, PR Metro Area	2011
B13015	Women 15 to 50 years who had a birth in the past 12 months by marital status and receipt of public assistance income in the past 12 months	Butler County School District, Alabama	2011

Table 4: Examples of Table Suppression for the 2011 1-Year Data Release

3 Appendices

3.1 Appendix A: DRB Memo

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

MEMORANDUM FOR:	James B. Treat Chief, American Community Survey Office
From:	Laura V. Zayatz Chairperson, Disclosure Review Board
Subject:	Approval of the Disclosure Rules for American Community Survey (ACS) Standard Data Products for the 2012 Data Year

This memorandum, along with the attached American Community Survey (ACS) 2012 Data Year Disclosure Review Board (DRB) Table Package, provides the DRB approved disclosure rules that apply to ACS standard data products for the 2012 1-year, 2010-2012 3-year, and 2008-2012 5-year data releases. The ACS standard data products include detailed tables (both base and collapsed), selected population profiles and other derived products including data profiles, subject tables, ranking tables, geographic comparison tables and the 1-year comparison profiles.

The derived products, consisting of estimates extracted from the detailed tables, do not need approval from the DRB once the detailed tables are approved, and therefore they are not specifically discussed in this document.

The American Community Survey (ACS) 2012 Data Year Disclosure Review Board (DRB) Table Package contains:

- Detailed tables: There are two tabs, one at the table level and one at the cell level.
  - Table level: Detailed table titles (both base and collapsed), the number of independent cells in the tables, data years for which the table is published, the table type, the geographic restrictions for the table, and the specific disclosure rules in this document that apply to the table,
  - Cell level: For each detailed table, the line-by-line contents, line number, row description, and sourcing for the tables.

- Selected Population Profiles (SPPs) tables: There are two tabs, one at the table level and one at the cell level.
  - Table level: The SPP table IDs, tables, description of the key column used to determine table eligibility for a given geographic area,

o   Cell level: For each SPP, the contents of the selected population profile tables, including the filtering groups and the key columns.

A.   DRB Rules on ACS Standard Data Products

This section outlines the specific rules from the DRB which shall be applied to all ACS data products, including the 1-year, 3-year and 5-year data products.

Selected Population Profile Case Minimums

1.   There must be at least 50 unweighted cases in the geographic area in the specific population subgroup "key column". The key columns are denoted in the "SPP Tables" tab in the attached spreadsheet. When there is a column that is suppressed by this rule, to show the table for that area, complementary suppression must be performed on other columns so that the suppressed column cannot be derived via subtraction. (To streamline the process, ACS suppresses whole tables instead of preforming complementary suppression.)

2.   There is an exception for SPP table S1703, Selected Characteristics of People at Specified Levels of Poverty in the Past 12 Months. Unlike other SPPs, the percentages are based on the row (stub line) percentage rather than the column (sub-population) percentage. Therefore, the 50 unweighted cases rule does not apply to the sub-population columns, just the total column (number of people for whom poverty is determined).

Non-Residence Geography Case Minimums

1.   Tables involving a geographic area other than current place of residence (such as workplace tables, place of birth, residence 1 year ago) crossed with characteristics other than current place of residence must have at least 50 unweighted cases in the universe of the table over the period.

Quantiles (Medians, Quartiles, Quintiles, Top Five Percent etc.)

1.   All quantiles in the ACS program are calculated using interpolation. A distribution of the variable in question is created, and the quantile is interpolated based on that distribution. Any quantiles that match a particular respondent's reported value are a coincidence only.

2.   Quantiles are calculated using a linear method. The bins for the distributions used for quantiles vary by topic, and are not necessarily uniform. The bins for the income quantiles are of an interval of at most 2,500, which allows the ACS program to not have to use the Pareto method of interpolation.

Means and Aggregates

1.   Estimates in the form of means or aggregates, defined here as a sum of the values for each of the elements in the universe (e.g., the sum of the income of all households in a

given geographic area), must be based on either zero unweighted cases or three or more unweighted cases (the Rule of 3) in a geographic area to show the mean or aggregate for that area.

2. When there is a mean or aggregate in a table for a given geographic area that is suppressed by this rule, complementary suppression must be performed on other means or aggregates to show that area so that the suppressed mean or aggregate cannot be derived via subtraction. (In practice, the ACS does not do complementary suppression, instead suppressing the whole table.)

3. Estimates in aggregate tables involving figures in dollars are rounded to the nearest 100. Estimates in aggregate tables involving travel time to work are rounded to the nearest 5 minutes. Estimates of aggregate number of vehicles used in commuting by workers 16 years and over by sex is rounded to the nearest 5. Estimates of aggregate hours worked, aggregate number of rooms, and aggregate number of vehicles available by tenure are not rounded. For estimates rounded to the nearest 100, estimates of (-100, 100), not inclusive, are rounded to 0. For estimates rounded to the nearest 5, estimates of (-7.5, 0), not inclusive, are rounded to -4 and estimates of (0, 7.5), not inclusive, are rounded to 4.

4. Tables on the number of people in a household and average household and family size, although technically aggregates and means, are not subject to the rules for means and aggregates.

5. If an aggregate income table is suppressed due to a line having only one or two unweighted cases, then the per capita income table associated with that aggregate table must also be suppressed, even if the number of unweighted people included in the per capita calculation is three or more.

Tables Not Allowed for Block Groups

1. Tables with more than 100 independent lines (excluding total and subtotal lines) cannot be released for block groups and tribal block groups. If a table is iterated by a variable such as race/ethnicity or gender, the set of iterated tables should be considered as a single table. The iterated variable should be considered a dimension when counting the number of lines in the table to determine if the set of iterated tables can be released for block groups and tribal block groups.

2. Certain other tables will not be published for block groups and tribal block groups. They include:
    - Tables where the universe is restricted to the foreign born or a subset of the foreign born;
    - Tables containing estimates of or characteristics of non-citizens;
    - Tables containing characteristics of unmarried partners;
    - Tables containing estimates of or characteristics of people that were married, widowed, divorced, or became mothers within the last 12 months;
    - Tables containing characteristics of people living in Group Quarters;

- Tables containing detailed type of Group Quarters (the only categories that can be shown at the block group level are Institutional and Non-Institutional);
- Tables containing detailed language categories (categories that can be shown at the block group level are English, Spanish, Other Indo-European, Asian/Pacific Islander, and All Else);
- Tables containing specific type of disability disabled by race, or categories of number of disabilities other than 0, 1, or 2 or more.

Unweighted Tables

1. Tables of unweighted counts of people and housing units may only be shown for areas where there are no occupied housing units or three or more unweighted occupied housing units.

B.   Rules Imposed by the ACS Program on ACS Standard Data Products

The population thresholds and data filtering rules described below are self-imposed by the ACS program for data quality purposes, not disclosure purposes. However, as part of the annual ACS standard data products submission to the DRB, these rules are reviewed by DRB and have influenced the final ruling from DRB on ACS data products. If these rules were not in place, the DRB may impose other restrictions, and so they will not be changed without DRB approval.

Population Thresholds

1. All geographic areas published for the 1-year ACS must have at least 65,000 people, and all geographic areas published for the 3-year ACS must have at least 20,000 people. If an area is included in the Population Estimates Program, then that estimate is used for the determination; otherwise, it is the ACS estimate.

2. There are two exceptions to the population thresholds:
    (a) All Metropolitan Statistical Areas with at least 50,000 people may be published for the 1-year ACS. This only applies to summary level 310 (Metropolitan/Micropolitan Statistical Areas); it does not apply so summary levels 312 and 314, which are metropolitan areas crossed by other types of areas.
    (b) An area that was published in the previous release may be published if it falls five percent or less below the threshold. To clarify, if an area had 66,000 people in Year 1, and 64,000 in year 2, it would be published in year 2, because the population is still above 61,750 (five percent below 65,000). If in year 3, if it had a population of 63,000, it would be published because it was published in the previous year and is over 61,750. However, in year 4, if it fell to 61,000, it would not be published and would not be published until it reached the 65,000 mark again.

3. For Selected Population Profiles, along with the case minimums described in Section A, there are two thresholds to determine if we publish a profile for the area:
    (a) For the 1- and 3-year ACS, there must be at least 500,000 people in the

geographic area, and

(b) There must be at least 65,000 people (for 1-year), 20,000 people (for 3-year), and 7,000 people (for 5-year) in the specific population subgroup "key column". (There is an exception for 1- and 3- iterated SPPs - an area that was published in the previous release may be published if it falls five percent or less below the threshold.)

Data Quality Filtering

1. The data reliability filtering rules described below only apply to the ACS 1-year and 3-year products, but not the 5-year products.

2. For detailed tables, a table is filtered out if the median coefficient of variation (CV) is greater than 0.61. A cell with an estimate of 0 is considered to have a CV = 1. Only lines at the lowest level of detail of a table are included in the calculation of the median CV.

3. Ratio tables are derived from other detailed tables. They are filtered out if the numerator or denominator estimates from the detailed tables is filtered out.

4. Cells in data profiles, subject tables and ranking/GCT tables are filtered out if the data from detailed tables that is used to source the cell is filtered out. Note that means that, unlike detailed tables, there can be tables with some lines filtered out and some lines not filtered out.

5. SPPs are generated directly from microdata and are filtered in a similar manner to detailed tables, except that the decision to filter out cells or not is made by cell groups as opposed to the whole tables (those cell groups are shown in the attached table packet). However, if one half or more of the lines in the SPP for a geography are filtered out, the whole table is filtered out for that geography.

6. For derived measures (medians, aggregates, ratios, proportions, rates, etc.) with a standard error equal to zero, the cell shall be suppressed if the estimated weighted total for the universe is less than 3,000. The test is evaluated for individual cells, without the requirement for complimentary suppression, meaning that a table containing multiple derived measures may be partially available. (If the estimated weighted total for the universe is 3,000 or more, an alternative formula is used to assign a non-zero standard error and suppression is not required.)

Comparison Profiles

1. Comparison profiles for a geographic area are published only if the area was published in the immediate previous year.

2. If a comparison profile is published for a geographic area, columns before the immediate previous year are published only if the area was published in that year.

Attachment:
- 2012 Data Year American Community Survey Disclosure Review Board Table Package

3.2   **Appendix B: List of Geographic Restrictions by Table**

List of Geographic
Restrictions by Table.