shadac
State Health Access Data Assistance Center
Bridging the gap between research and policy

Robert Wood Johnson Foundation

September 2012

Understanding 1-, 3-, and 5-year ACS Estimates: Summary Tabulations and Public Use Files

Introduction

The American Community Survey (ACS), conducted by the U.S. Census Bureau, is an ongoing general household survey of the entire population. This survey replaced the long form of the decennial census and provides annual estimates of economic, social, demographic, and housing information for the nation, states, and sub-state geographies. The ACS became fully operational in 2005 and health insurance coverage questions were added in 2008.

The Census Bureau provides three ACS data products: 1-, 3-, and 5-year estimates in published tables and in public use files. The purpose of this document is to explain the differences between these data products, provide guidance on when to use single-year vs. multi-year estimates, and provide some technical details about the multi-year files.

Published Tables (Summary Tabulations)

The Census Bureau publishes 1-year estimates for areas with populations of 65,000 or more; 3-year estimates (covering 36 months) for areas with populations of 20,000 or more; and 5-year estimates (covering 60 months) for all statistical, legal, and administrative entities. Multiple years are pooled together to produce reliable period estimates for areas and subgroups with smaller population; they are not a simple average of the years covered. Table 1 describes the level of geography available for the 1-, 3-, and 5-year estimates. About 25% of counties are available from the 1-year estimates, about 59% from the 3-year estimates, and all U.S. counties are available from the 5-year estimates. Published estimates for the geographies described in Table 1 are available from the Census Bureau's American FactFinder (AFF) at http://factfinder2.census.gov.

Table 2 lists the years of data availability. The first release of the 3-year health insurance coverage estimates was in 2011, based on the 2008-2010 surveys, and the first release of 5-year health insurance coverage estimates will be in 2013, based on the 2008-2012 surveys.

Table 1. Geography for which ACS estimates are published

	1-year	3-year	5-year
Published Geography [#]	Areas with populations > 65,000	Areas with populations > 20,000	All areas
Nation, States, and Dist. of Columbia	All	All	All
Congressional Districts	All	All	All
Counties	About 25%	About 59%	All
School Districts	About 7%	About 24%	All
Census Tracts	None	None	All

[#]For a comprehensive list of geography available for the ACS 1-year, 3-year, and 5-year estimates see http://www.census.gov/acs/www/guidance_for_data_users/geography/

Table 2. Years for which ACS estimates are published

1-year	3-year	5-year
2005		
2006		
2007	**2005-2007**	
2008	**2006-2008**	
2009	**2007-2009**	**2005-2009**
2010	**2008-2010**	**2006-2010**
2011 (Available Sep. 2012)	**2009-2011** (Available Oct. 2012)	**2007-2011** (Available Dec. 2012)
2012 (Available 2013)	**2010-2012** (Available 2013)	**2008-2012** (Available 2013)

Shading indicates years with health insurance coverage available

Public Use Microdata Sample

In addition to publishing summary estimates, the Census Bureau provides Public Use Microdata Sample (PUMS) files. The PUMS files are a subset of the full file and a single-year represents about 1% of the U.S. population and contains about 3 million individual records. The lowest level of geography available for the 1-, 3-, and 5- year PUMS is the Public Use Microdata Area (PUMA). PUMAs are designed to disseminate the data while meeting disclosure requirements so individuals cannot be identified. They are made up of areas of about 100,000 people, are contained within states, and are most often groups of counties or parts of counties.

PUMS files allow analysts to create custom tabulations for estimates that aren't available from the summary tabulations. Analytic software such as SAS, Stata or SPSS is needed to process custom tabulations. SHADAC tabulates health insurance coverage estimates from the PUMS files for categories relevant to health policy. For example, SHADAC tabulations include health insurance coverage estimates by Affordable Care Act (ACA) policy-relevant groups including young adults through age 25 and people living at or below 138% of federal poverty guidelines.[1] These custom estimates are available from SHADAC's Data Center, an easy to use online table generator, at http://www.shadac.org/datacenter.

[1] Under the ACA adults are allowed to stay on their parent's health care plan through age 25. At or below 138% of federal poverty guidelines are the new national income eligibility guidelines for most non-elderly adults under 2014 Medicaid expansions.

The PUMS files are available from AFF and also from the University of Minnesota's Integrated Public Use Microdata Series (IPUMS) at http://usa.ipums.org/usa. IPUMS harmonizes the variable names and labels over time allowing for easier trend analysis, and allows you to select only the variables you want to analyze. IPUMS also provides not only data, but codebooks and command files for SAS, Stata and SPSS. IPUMS also has SHADAC's health policy-relevant variables for the health insurance unit and federal poverty guidelines (SHADAC 2012).

Single-Year vs. Multi-Year Estimates

As more years are pooled together the estimates become less current, but the statistical reliability of the estimates increases. Single-year files provide the most current data, so these are a good choice if the level of geography of interest is available and there is a large enough sample size for any sub-group analysis of interest. Multi-year files are less current, but with a larger sample size are more reliable. They allow analysis of smaller geographies and smaller sub-groups.

While the lowest level of geography available on the PUMS files for the 1-, 3-, and 5-year files is the same, the PUMA, using multi-year data vs. single-year data allows for analysis of demographic subpopulations within a geography. The more you refine your subpopulation (for example, children living in poverty by race/ethnicity and by state) the more likely you will need to use multi-year data to have enough sample size for reliable estimates.

Table 3 provides an example of estimates available from the 1-, 3-, and 5-year summary tabulations for a subgroup with a small population in Wyoming. Estimates are available for 2 counties in Wyoming using the 1-year data, 10 counties using the 3-year data, and all 23 counties using the 5-year data. The estimates for the American Indian and Alaska Native population are more reliable in the 3-year and 5-year data due to the larger sample size. The 1-year estimate for Natrona County is 4.2% with a margin of error of 1.5%, the 3-year estimate is 2.4% with a margin of error of 0.6%, and the 5-year estimate is 2.3% with the smallest margin of error of 0.3%.

Table 3. American Indian and Alaska Native Population for Wyoming Counties: ACS 1-year, 3-year, and 5-year Estimates

	1-year 2010		3-year 2008-2010		5-year 2006-2010	
County	Percent	CI	Percent	CI	Percent	CI
Albany			1.8	±0.4	1.7	±0.2
Big Horn					3.3	±1.7
Campbell			2.3	±0.7	2.1	±0.4
Carbon					1.9	±0.4
Converse					2.2	±0.8
Crook					0.3	±0.2
Fremont			23.3	±0.8	22.9	±0.6
Goshen					2.0	±0.8
Hot Springs					1.9	±1.0
Johnson					1.3	±1.2
Laramie	2.2	±0.7	2.1	±0.4	2.0	±0.3
Lincoln					1.2	±0.5
Natrona	4.2	±1.5	2.4	±0.6	2.3	±0.3
Niobrara					4.2	±2.5
Park			1.6	±0.6	1.8	±0.6
Platte					0.8	±0.8
Sheridan			2.2	±0.8	2.4	±0.2
Sublette					11.0	±4.9
Sweetwater			4.0	±1.0	3.3	±0.6
Teton			2.0	±1.3	1.7	±0.9
Uinta			4.4	±1.2	3.8	±1.0
Washakie					2.0	±1.0
Weston					2.0	±0.8

Source: U.S. Census Bureau 2010, 2008-2010, and 2006-2010 American Community Survey's (http://factfinder2.census.gov), accessed September 6, 2012. CI (confidence interval) can be interpreted roughly as providing a 90 percent probability that the interval defined by the estimate minus the CI and the estimate plus the CI (the lower and upper confidence bounds) contains the true value. The American Indian and Alaska Native population is race alone or in combination with one or more other races.

Comparing Estimates

The Census Bureau recommends that comparisons only be made within data products:

- 1-year estimates with 1-year estimates
- 3-year estimates with 3-year estimates (non-overlapping years)
- 5-year estimates with 5-year estimates (non-overlapping years)

They advise against comparisons across single-year and multi-year files. For example, they advise against comparing 1-year state estimates with 5-year county estimates. If interested in this comparison you should use both state and county estimates from the 5-year file.

The Census Bureau also recommends that comparisons with the multi-year files be made with non-overlapping years. For example, 2005-2007 can be compared to 2008-2010, but not 2007-2009. Making comparisons and performing statistical testing is easier to interpret if overlapping years are not included. Guidance and formulas for making comparisons across time periods are available in Census Bureau documentation (U.S. Census Bureau 2009a).

Technical Details for Multi-Year Files

While not comprehensive, this section contains technical details about the creation of the multi-year files that are important to understand when analyzing the data. The Census Bureau provides some helpful presentations that provide guidance on using the ACS (U.S. Census Bureau 2009b).

Weighting

The multi-year year files contain the estimates from each of the single years included in the file, but the estimates are reweighted to provide a pooled estimate covering the entire 3- or 5-year time period (U.S. Census Bureau 2011a). The pooled estimates are not a simple average of the years included in the time period, so while rare, it is possible for the multi-year estimate to fall outside the range of the single years included.

Population Controls

As part of the weighting procedure, the estimates are controlled to population estimates from the Census Bureau's Population Estimates Program. The 1-year 2008 and 2009 estimates use population estimates based on the 2000 Census updated for births, deaths, and migration (postcensal). The 1-year 2010 estimates use population estimates based on the 2010 Census. Because of the change in population controls, the Census Bureau recommends that data users exercise caution when comparing 2009 and earlier with 2010 single-year estimates (U.S. Census Bureau 2011b).

In the multi-year files the estimates are controlled to the current version of population estimates. The 3-year file (2008-2010) estimates are controlled to population estimates that have been modified to account for the 2010 Census (intercensal). Because of this, a consistent methodology is used, so caution is not needed when spanning a decennial census in a multi-year file (U.S. Census Bureau 2011c).

Income and Geography Adjustments

Income is inflation adjusted, using the Consumer Price Index, to represent the final year in the multi-year period. Geography may change over time, and the multi-year estimates are based on the geography of the final year in the period. For example, say estimates were published for a large city in 2008 and 2009. In 2010 this city annexed surrounding territory, changing the geographic boundaries for the 2010 estimates. The 3-year file for 2008-2010 uses the 2010 geographic boundaries.

Logical Edits for Health Insurance Coverage

Beginning in 2009 the Census Bureau added a series of logical edits to the editing routine for health insurance coverage. These logical edits are intended to improve the reporting of insurance coverage in the survey. Logical edits assign coverage to Medicare, Medicaid, and TRICARE/military coverage based on the respondent's answers to other items in the questionnaire (SHADAC 2010). For example, a person who is active duty military, and has not selected TRICARE/military coverage, is automatically assigned that coverage. In 2008, logical edits reduced the uninsurance rate from 15.1% to 14.6%.

The Census Bureau has not updated single-year 2008 estimates in AFF or the PUMS file with the logical edits, but the logical edits have been added to 2008 in the 3-year file and will be included in the 5-year file when it is released in 2013. SHADAC has applied the logical edits to the single-year 2008 PUMS file and made it available from IPUMS.

Suggested Citation

State Health Access Data Assistance Center. 2012. "Understanding 1-, 3-, and 5-year ACS Estimates: Summary Tabulations and Public Use Files." SHADAC Brief #32. Minneapolis, MN: University of Minnesota.

About SHADAC

The University of Minnesota's State Health Access Data Assistance Center (SHADAC) is funded by the Robert Wood Johnson Foundation to collect and analyze data to inform state health policy decisions relating to health insurance coverage and access to care. For information on how SHADAC can assist your state with data issues relevant to state health policy, please contact us at shadac@umn.edu or call 612-624-4802.

References

State Health Access Data Assistance Center. 2010. "A Summary of the American Community Survey Logical Edits Applied to Health Insurance Coverage." Minneapolis, MN: University of Minnesota. Available at http://www.shadac.org/files/shadac/publications/TechBrief_ACSEdits_Dec2010.pdf

State Health Access Data Assistance Center. 2012. "Defining 'Family' for Studies of Health Insurance Coverage." Issue Brief #27. Minneapolis, MN: University of Minnesota. Available at http://www.shadac.org/publications/defining-family-studies-health-insurance-coverage

U.S. Census Bureau. 2009a. "A compass for Understanding and Using American Community Survey Data What Researchers Need to Know." Available at http://www.census.gov/acs/www/Downloads/handbooks/ACSResearch.pdf.

U.S. Census Bureau. 2009b. American Community Survey Training Presentations web page. Available at http://www.census.gov/acs/www/guidance_for_data_users/training_presentations.

U.S. Census Bureau. 2011a. "American Community Survey Multiyear Accuracy of the Data." Available at http://www.census.gov/acs/www/Downloads/data_documentation/Accuracy/MultiyearACSAccuracyofData2010.pdf.

U.S. Census Bureau. 2011b. "American Community Survey Research Note Change in Population Controls." Available at http://www.census.gov/acs/www/Downloads/comparing_acs_data/2010_Change_Population_Controls.pdf

U.S. Census Bureau. 2011c. "Methodology for the Intercensal Population Estimates: 2000 to 2010" available at http://www.census.gov/popest/methodology/2000-2010_Intercensal_Estimates_Methodology.pdf.