How to Use the American Community Survey (ACS)

Guidance on Application of American Community Survey (ACS) Data

For information on the applications of American Community Survey data, including the use of thresholds for determining reliability, see A Compass for Understanding and Using American Community Survey Data: What State and Local Governments Need to Know.

On the Use of Estimates from a Continuous Survey

Compared to decennial census data, ACS estimates require a different frame of reference since they are derived from a continuous survey. In the past, data users considered census estimates as being "pinned" on a specific reference point: April 1 of the census year. The ACS 1-year estimates provide a picture that represents an average over a 12 month period, 3-year estimates represent averages over 36 months, and 5-year estimates are 60 month averages. The Census Bureau refers to these estimates as "period estimates," since they represent the best estimate for a characteristic over a specified period of time: one, three or five years (e.g. the population of Queens was 48 percent foreign-born in the 2006-2008 period).

On Choosing an Appropriate ACS Data Source

Choosing the right ACS data source is largely a matter of balancing the following three data characteristics according to one's research needs:

1. ***Geography –*** what level of geography is needed to address the question at hand?

2. ***Timeliness of the data*** – is it essential that you have data for the latest time period?

3. ***Reliability –*** how statistically reliable do estimates need to be for the application at hand?

Achieving the best compromise is closely related to the data application. Some users may decide that timeliness is the highest priority and will sacrifice geographic detail, opting to use the single year data for an entire borough, instead of 3-year data for a PUMA of interest. For example, if a recent change in the economy is best reflected in the data for 2013, the borough number may better reflect this downturn than a 3-year PUMA estimate that includes data for 2011 and 2012.

For others, geography will trump timeliness, making it more acceptable to use a 3-year period estimate for a single PUMA than the 1-year estimate for the borough. Such is the case when attempting to demonstrate a characteristic that is unique to a PUMA, such as the number of foreign-born persons with limited English proficiency (LEP); targeting programs requires geographic detail below the borough level. Similarly, there are cases where needs exist within geographic areas that may not be apparent at the borough level, as when a government is trying to determine whether elderly homeowners are in need of assistance; multiyear PUMA estimates should work well here.

Finally, there may be cases where reliability is less important than geography and timeliness. This is the case where a general idea of a characteristic is sufficient to answer a question, as with the presence of foreign-born population or a specific foreign-born subgroup. Knowing the precise percentage of population foreign-born or the precise number from a particular subgroup may not be important; getting a general idea of those areas that are at both ends of a continuum may be the goal.

For large geographies, such as New York City and the five boroughs, the option exists for reliable 3-year and 1-year estimates. Because most of these estimates are reliable, users will likely opt for the 1-year data since they are more timely than 5-year estimates. There may be situations, however, where small subgroups of the population (e.g. new immigrant groups) could be more difficult to estimate reliably with one year of data, compared with estimates using data for three years. As we proceed with new ACS data each year, the choices for comparisons will expand. The biggest change occurred in 2010-2011, when the first five-year estimates were released for census tracts, PUMAs, boroughs, and the city. Despite the addition of five-year period estimates the choice of ACS datasets is still related to the question at hand and the priority of the three items above.

Regarding the reliability of estimates, it should be noted that many data users, including the Census Bureau, are now using c**oefficients of variation (CVs)** to better understand the reliability of ACS estimates. CVs put an estimate's margin of error in relation to the estimate itself; almost as if the error term were a percentage of the estimate. Lower CV values suggest a more reliable estimate. There is no standard endorsed CV threshold, because acceptable CVs depend on how ACS data are being applied. However, users should keep in mind that a CV of 20 means that the corresponding estimate may vary, plus or minus, by one-third within the 90% confidence interval. For a median household income of $60,000, that would mean the estimate could reasonably be as high as $80,000 or as low as $40,000; a degree of variability that may be unacceptable for many applications. For more information on CVs please refer to page 5, or Appendix 3, in *What State and Local Governments Need to Know* in the Compass Series publications.

On Comparing ACS Estimates

The following lists the ways users should exercise caution when making comparisons that involve ACS estimates.

Compare period estimates of the same duration: 1-year to 1-year; 3-year to 3-year, 5-year to 5-year. Comparisons across geographies should employ period estimates of the same duration. For example, a comparison of the median household income between Los Angeles and New York City could compare a 1-year estimate to another 1-year estimate (e.g. 2013 vs. 2013 for each city), but should not compare a 1-year estimate with a 3-year estimate (e.g. 2012 vs. 2011-2013). Comparisons across time should also employ period estimates of the same duration. For example, an examination of the change in Brooklyn's unemployment rate could analyze the difference between the rate in 2008-2010 and the rate in 2011-2013. However, it would not be appropriate to conduct the same comparison looking at the change between the 2008-2010 period estimate and the 2013 period estimate.

Confirm that differences are statistically meaningful by using significance testing. Because ACS estimates are derived from a sample, all differences are not necessarily statistically

significant. In order to confirm that differences between estimates are statistically significant (i.e., the differences are meaningful according to sampling theory) users should apply a significance test. Guidance on this topic can be found in Appendix 4 of the Compass Series publications.

Exercise caution when examining change across time using estimates that have overlapping years. The Census Bureau advices against comparisons across time involving overlapping years (e.g. 2009-2011 vs. 2011-2013). However, there are instances when such comparisons are unavoidable. When conducting comparisons with overlapping years be sure to test that the difference between the two estimates are statistically significant. Guidance on this topic can be found in Appendix 4 of the Compass Series publications.

You can compare 2000 Census data with data from the ACS, but approach these comparisons with caution. While the ACS is a replacement for the decennial census long form, it is NOT the long form. Differences in residence rules, reference periods for items like income and previous residence, and changes in how questions are asked make interpretation of differences problematic. Further, since the ACS is a continuous survey, changes occur from year to year in the sample, content, and conduct of the survey. Users need to know about these changes in order to exercise caution, where appropriate. Detailed information on the comparability of different years of the ACS and of the comparability of the ACS with the 2000 Census are available at:
http://www.census.gov/acs/www/guidance_for_data_users/comparing_data/

Special Note about Multi-year Estimates

Users of multi-year population estimates should recognize that these estimates (e.g., the number of persons in poverty) are average counts across a number of years. Some areas may have experienced appreciable increases or decreases over a 3- or 5-year period (e.g., in the number of high school graduates, or persons below the poverty line), making it difficult to interpret what an average for the period actually represents. For example, a 3-year period estimate that 5,000 people in a PUMA live in poverty could reflect any of the following: a constant 5,000 in each of the three years, a steady increase from 3,000 to 5,000 to 7,000 persons; a corresponding steady decrease (7,000 to 5,000 to 3,000); or even a rise and decline in the percentage across the years, say 3,000 to 8,000 to 4,000; and so on. This means that multi-year averages for all variables will "flatten-out" underlying patterns that may be present in the data. To obtain an indication of the likely pattern that underlies a 3-year estimate, users need to apply local knowledge of the conditions in an area over the period, such as shifts in economic cycles. For example, the onset of a recession may cause one to conclude that there was an increase in the number of persons below the poverty line over a three year period in a PUMA. Or the median value of owner-occupied housing units in a PUMA could vary greatly over the course of a three year period as a result of economic conditions. This means that the PUMA's 3-year average may not be a useful current indicator of its status for program planning or resource allocation.