The American Community Survey: Tracking How We Change With Multi-Year Estimates

Cynthia M. Taeuber (cmtaeuber@comcast.com)
Association of Public Data Users, November 18, 2009

Data sources for local areas:

Four data sets from the Census Bureau – 2010 Census counts, the American Community Survey (ACS), the Local Employment Dynamics (LED) program (http://lehd.dsd.census.gov/led/) , and the Population Estimates Program (PEP) (http://www.census.gov/popest/estimates.html – all help local areas understand and compare their population, economic, and housing characteristics. The ACS gives a broad picture of the characteristics of the household population and housing of areas, while the LED tells us about the characteristics of workers in the industries and areas where they work and live. Between censuses, PEP gives us estimates of the number of people with limited demographic characteristics. Some administrative records from programs are available for research purposes as well.

American Community Survey Basics:

The ACS (http://www.census.gov/acs/www/index.html) is not your mother's decennial census. It produces annual and updated multi-year estimates of the CHARACTERISTICS of the population and housing of areas for the same geography and topics as the decennial census long form. Counts of the number of people and housing units will continue to come from decennial censuses and from PEP between censuses. The ACS website shows the questionnaires and products and describes the survey's methodology. You can find ACS data on the Census Bureau's American FactFinder (http://factfinder.census.gov). Also see: http://www.census.gov/acs/www/Products/

Starting in 2010, **three estimates** will be produced for areas with 65,000+ people. There will be a 1-year estimate from 2009, a 3-year estimate from 2007-2009, and a 5-year estimate from 2005-2009. Once an area receives a 3-year or 5-year estimate, that area will receive an annually updated multi-year estimates each year thereafter. Five-year averages for the smallest areas, including census tracts and Block Groups,* will start with the data for 2010 (released in 2011). The Census Bureau will not publish estimates for small numbers of people or areas if there is a probability that an individual can be identified.

Data Availability

--Areas and groups of 65,000 or more people will have 1-, 3-, and 5-year estimates.

--Areas and groups of 20,000 to 64,999 people will have 3- and 5-year estimates.

--Areas and groups of fewer than 20,000 people will have 5-year estimates only.

*Census tracts are small, relatively permanent statistical subdivisions of a county averaging about 4,000 inhabitants. Census Block Groups (BG) generally contain between 600 and 3,000 people and are the smallest geographic level for which data will be produced. BGs are not intended as individual units of analysis because sampling variability for individual BGs is generally too large to make comparisons or statements about change. We recommend that Block Groups be added together to create reasonable sample sizes and reliable estimates.

For 2000 – 2005, data are for the household population only. Data for the total population first became available in 2006 when the group quarters population was added to the survey.

Reference date: The ACS single-year estimate is for the characteristic averaged over 12 months of a calendar year; for a 5-year estimate, it is over 60 months. Each month contributes equally. It is ***not*** an estimate of an implied beginning, mid- (such as July 1), or end-of-year characteristic.

The effect of sampling and sampling error on the analysis:

The sample size for the decennial census long form was about 17 percent of households. The BASIC statistic from the ACS for all areas, regardless of size, that is most comparable to the long form is the 5-year average although the sampling rate is smaller. Currently, annual sample is about 3 million housing units and the initial sampling rate has eroded from about 2.4 percent of households in 2005 to 2.1 percent in 2009.

The *smaller* the sample size, the *larger* the variation in the estimate. The ACS publishes the margin of error (MOE) of the estimate. **Pay attention to the variation or you will make stupid mistakes**. What appears to be change across a year or two may be nothing more than the variation that comes from sampling. In your analyses, consider whether differences in the estimates are statistically significant as well as of any practical importance. For information about the multi-year estimates, see: http://www.census.gov/acs/www/AdvMeth/Multi_Year_Estimates/overview.html

Show me the data

In write-ups of your analyses, most of the time, you will use percentage distributions in the text rather than numbers because they are more meaningful for comparisons. To learn how to compute whether apparent differences are statistically significant, see "Accuracy of the Data" on the ACS website (http://www.census.gov/acs/www/UseData/Accuracy/Accuracy1.htm).

A vexing question is how to incorporate information about statistical uncertainty into what you write without overburdening the reader with numbers. There are many options, from the use of the word "about," to putting the margin of error or the confidence interval in parentheses after each number.

How do you compare geographic areas that have different multi-year averages?

It depends on your purpose. If you need:
Consistency – Use the 5-year averages for all areas (or 3-year averages for areas of 20,000+ people)
Replace long form estimates – The 5-year estimates replace the decennial census long-form statistics
Reduce lag time – For larger areas, the 3-year estimates have less lag and reasonable reliability
Stability and reliability – How sure do you need to be? There is a tradeoff between the *timeliness* of the 1-year estimates (finger on the pulse but larger confidence interval), the more stable 3-year estimates, and the *lag* of the trailing indicators from the 5-year moving averages.

In comparing 5-year averages, compare independent periods with no overlap (e.g., 2005 – 2009 compared with 2010 – 2014) to monitor trends in a characteristic. In a comparison of, for example, 2005 – 2009 and 2006 – 2010, four of the 5 years are the same estimate and highly correlated and really, you are comparing only the single-year estimates of 2005 and 2010. For smaller areas, even though you cannot measure year-to-year differences precisely, you can monitor long-term trends, the magnitude and direction of the trend, and the relative differences among areas or population groups, from the *series* of annually updated estimates. Use the 1- and 3-year averages to detect sudden and large changes.

For more information:

American Community Survey website: www.census.gov/acs/www/ (subscribe to the ACS Alert for free periodic updates). Call the ACS staff: **1-800-923-8282**. The Census Bureau has developed a series of educational materials called *The ACS Compass Products*:
http://www.census.gov/acs/www/UseData/Compass/compass_series.html

Acknowledgement: Discussions with Census Bureau statisticians David Hubble and Alfredo Navarro and a paper presented to the American Statistical Association by Charles Alexander, Jr., contributed to the development of this information.