Are You Ready for the American Community Survey?

Liesl Eathington
Department of Economics
Iowa State University

December 13, 2010

Data that describe characteristics of the population, whether used for academic research and teaching, government policy and program development, business planning and marketing, media stories or myriad other applications, are in high demand. That is why socio-economic data users of all stripes may be interested to know that on December 14, 2010, the U.S. Census Bureau plans a major release of data from its American Community Survey (ACS).

If you ever use statistics about socio-economic characteristics of the population, then you are a potential user of ACS data. In fact, you may already be using ACS data without knowing it. The Census Bureau has been developing the ACS for more than a decade and has been gradually rolling out data since 2001. Data have been available for geographic areas with 65,000 or greater population since 2006, and for areas with 20,000 or greater population since 2008. With this month's release, the U.S. Census Bureau reaches an important milestone in the development of the ACS with its first release of data for smaller geographic areas.

Many people are still unaware of the ACS, despite its long development period and the fact that state-level ACS data have been accessible for several years. Others, who are familiar and comfortable with traditional decennial Census data, are somewhat leery of the changes that will be required in adapting to the new ACS framework.

The next section of this paper provides a brief overview of the American Community Survey (ACS). Readers who are already familiar with the ACS may wish to skip directly to the quiz in the subsequent section in order to test their readiness for using the 5-year estimates. The last section describes where to find ACS data and where to find guidance in using ACS data.

Brief Description of the ACS

The ACS is a large, continuous survey that samples approximately 3 million addresses in the United States per year. The survey produces estimates of population and housing characteristics for specific geographic areas and population groups. Questions on the ACS collect information about:

- age and sex
- race and ethnicity
- family and relationships
- income and benefits
- health insurance
- education

- veteran status
- disabilities
- work status, where you work, and how you get there
- characteristics of your residence and housing costs

In all, the upcoming release of 5-year ACS data will provide estimates on 72 social, economic, housing and demographic topics for all communities, counties, and states. Of particular interest to researchers, data will be available down to a very high level of geographic detail. The distinct geographic areas described in the upcoming release will include census block groups, census tracts, townships and other sub-county areas, school districts, counties, and states.

Producing such detailed characteristics of a particular geographic area requires a large sample size to ensure reliability. One year of data collection is sufficient for entire states and other areas with a large population. Multiple years may be required to obtain a sufficient sample in less populous areas. Pooling the results from multiple survey years allows the Census Bureau to produce reliable estimates for these smaller areas. Under the ACS framework, the Census Bureau will release new data sets annually for all geographic areas based on a 1-year, 3-year, or 5-year sample, depending on the size of the area:

- 1-year data sets are available for areas with a population of 65,000 or more. In Iowa, there are currently six cities and eight counties for which 1-year data sets are produced. The first 1-year data sets, which described 2005, were released in 2006. New 1-year data sets have been produced annually since then. The most recent 1-year data sets, which describe 2009, were released in September 2010.
- 3-year data sets are produced for areas with a population of 20,000 or more. Iowa has 22 cities and 35 counties for which 3-year data are produced. The first 3-year data sets, with data for 2005-2007, were released in 2008. The most recent 3-year data sets cover the period from 2006-2008, and data for 2007-2009 are scheduled for release in January 2011.
- All other geographic areas will have their first glimpse of ACS data with the December 2010 release of 5-data. Those 5-year estimates are based on survey responses spanning the five years from January 1, 2005 and December 31, 2009. Each year, the Census Bureau will release a new set of 5-year estimates covering the most recent five-year period.

American Community Survey Quiz

Many members of the media, the academic research community, policy-makers, and educators have eagerly anticipated the U.S. Census Bureau's release of the first set of 5-year estimates from the ACS. The following set of true/false questions test your knowledge about some of the features of the 5-year data and the ACS in general.

Questions

1. *True or False*: The new ACS data release is based on information collected during the recently-completed 2010 Census.

2. *True or False*: The American Community Survey has replaced the long form instrument that used to be part of the decennial census.

3. *True or False*: Because participation in the ACS is voluntary, ACS data will be subject to greater sampling bias than data from the Decennial Census.

4. *True or False*: Data for small cities will now be available from the Census every 5 years instead of every 10 years.

5. *True or False*: If I want to compare per capita income across cities in Iowa, I should use the most current data available for each place; for example, the 1-year ACS data for Des Moines (2009), the 3-year estimate for Fort Dodge (2007-2009), and the 5-year estimate for Carroll (2005-2009).

6. *True or False*: I can compare the 5-year ACS data with the 2000 Census data to measure how Iowa's residents have fared during the recent recession.

7. *True or False*: Because the 5-year estimates are most reflective of the midpoint of the time period, the new 2005-2009 estimates best describe economic conditions at the beginning of the recent recession.

8. *True or False*: I can ignore the margin of error estimates because they are only important to egg-heads.

9. *True or False*: The 5-year ACS estimates that will be released next year (covering 2006-2010) will be more accurate than the 2005-2009 data released this year.

10. *True or False*: The dollar values reported in the 5-year ACS data have already been adjusted for inflation.

Answers

1. FALSE. The American Community Survey is an ongoing data collection effort that is distinct from the comprehensive population count conducted for the 2010 Census. The upcoming release of 5-year ACS data is based on survey responses spanning the five years between January 1, 2005 and December 31, 2009.

2. TRUE. The so-called "long form," historically sent to a subset of households during the decennial census effort, was not included as part of the 2010 Census. The 2010 Census included only a short, 10-question instrument. Researchers and others who have relied on detailed data from the long form in the past may encounter some challenges in adapting to this change, particularly when trying to compare ACS data with previous census results for analysis of changes over time. Researchers are encouraged to seek additional guidance on this topic from the U.S. Census Bureau.

3. FALSE. Participation in the ACS is not voluntary, it is mandatory. The Census Bureau follows up on non-responders using similar methods that they employ for the decennial census. In some ways, ACS data may be slightly superior in quality than decennial census data because the survey is administered by professional Census Bureau employees rather than temporary employees used for the decennial count.

4. FALSE. New data will be available EVERY year for all communities, regardless of size. However, for small communities, each year's new data release will represent a rolling survey from the previous 5 years. For slightly larger communities, the new data will represent a rolling survey from the previous 3 years. For the largest communities, the new data will represent an annual survey. The

Census Bureau anticipates that the lag time for new data releases will be much shorter than we've been accustomed to for decennial census data products. The Census Bureau has been releasing ACS data within a year of completing the data collection.

5. FALSE. When comparing estimates across different geographies or different population groups, you should always use estimates from data sets covering the same time frame. That will require using the data set for your smallest place or population group. In the example above, your analysis should use only the 5-year estimates for all of the cities you are comparing.

6. FALSE. The data in the 5-year release were collected continuously from January, 2005 through December, 2009. Because the 5-year data release pools the responses from before, during, and after the recession, it cannot provide an accurate reflection of how the recession has affected Iowa's households. In addition, the Census Bureau has advised a general caution for comparing ACS results with 2000 Census results. Slight changes in content and wording of questions will make some comparisons more difficult than others. Users are encouraged to look for more information about this topic in the Census Bureau's extensive library of ACS-related publications.

7. FALSE. The multi-year estimates are no more representative of the middle year than the beginning or ending year of the measurement period. The estimates are derived from the mean of the distribution of all values collected during the entire period. How closely the estimated value aligns with the actual value in any particular year depends on the underlying behavior of the variable in question, which could be increasing, decreasing, fluctuating, or flat.

8. FALSE. The margins of error associated with the ACS estimates are extremely important for comparing values across regions or changes over time. The estimates are based on a survey, which introduces some degree of uncertainty about how well the estimates describe the underlying population. Historically, data produced from the long form of the decennial census were also based on a sample; however, margins of error associated with the long form data were rarely published and generally ignored outside the research community. The Census Bureau is strongly encouraging data users to acknowledge and accept the uncertainty associated with the sample data and to make use of the margins of error when comparing estimates. Data users, especially those with an aversion to statistics, will likely resist mightily, creating a great need for education and technical assistance from the Census Bureau, its affiliated State Data Centers, and the academic community.

9. TRUE (qualified). The 2005-2009 ACS numbers released this month have not been benchmarked with the results of complete count of the population that was conducted as part of the 2010 Census. Consequently, measures that are based on underlying assumptions about the distribution of our population by age, race, ethnicity, household size, etc. are still based on 2000 Census data. We will have more confidence in next year's ACS data release, especially for small areas and population groups, after the 2010 Census counts have been incorporated into the estimates. That said, the same underlying sampling and estimation methodologies will be used in producing the pre-census and post-census estimates.

10. TRUE. For its multi-year estimates, the Census Bureau adjusts financial values so that they are consistent across the whole data collection period. For example, the estimates in the 5-year data release for 2005-2009 have all been adjusted and restated in values equivalent to 2009 dollars.

Sources for ACS Data and Assistance

U.S. Census Bureau (http://www.census.gov/acs/www/)

The U.S. Census Bureau, which produces the ACS, is the primary source for obtaining ACS data as well as guidance on using and interpreting ACS data. The Census Bureau has devoted an entire section of its Web site to the ACS. Users can download entire ACS data sets and documentation from the ACS library, located at the URL listed above.

Users can also access information for a particular community or groups of communities by using the Census Bureau's American FactFinder utility (http://factfinder.census.gov/home/saff/main.html). While it may require some practice before users appreciate the functionality of American FactFinder, the time is well worth the effort. American FactFinder provides data from a number of other Census Bureau surveys and programs, including the 1990 and 2000 Census.

In addition to the data themselves, the Census Bureau has provided an extensive library to assist users in understanding and using ACS data. That library includes a treasure trove of user guides for the media, teachers, and researchers, multi-media presentations, and other resources.

State Data Center Network of Iowa (http://www.iowadatacenter.org/)

Through its network of State Data Centers (SDCs), the U.S. Census Bureau seeks to disseminate both data and technical assistance to a wide audience of users. The lead agency for the State Data Center of Iowa is housed in the State Library of Iowa. The Iowa SDC operates its own Web site, located at the URL referenced above. In coming months, the 5-year ACS data will be made available at this site. Currently, the site provides the 1-year (from 2005 onward) and 3-year (2005-2007 onward) ACS data sets as well as numerous other socio-economic data for Iowa's school districts, cities, counties, and other areas.

As a coordinating agency within the State Data Center of Iowa's network, Iowa State University will also provide researchers and the broad data-using public with assistance in using and interpreting ACS data. In addition, Iowa State University will be issuing ACS-related data products and reports in coming months. Watch the Web site for the Regional Economics and Community Analysis Program (http://www.recap.iastate.edu) for more information and updates about this valuable new source of data.

Other Sources

The 5-year ACS data will very soon be incorporated into countless other information sources, including many Web-based socio-economic data applications. It is likely that not all of these sites will provide the supporting documentation necessary to correctly interpret the ACS data. Where technical documentation is lacking, users are encouraged to consult information provided by the U.S. Census Bureau or affiliated institutions and agencies.

Summary

This paper is intended as a brief introduction the American Community Survey and to test your readiness for using new ACS data. It introduces some key issues, addresses some common misconceptions, and raises some warnings about using ACS data.

Given all of the cautions and limitations surrounding use of the ACS data, many people are wondering, "Just what CAN we use these data for?" In general, the Census Bureau provides this advice:

- Use numbers from the 2010 Census to obtain counts of the population and their basic characteristics (sex, age, race, Hispanic origin, and homeowner status).
- Use data from the American Community Survey to obtain demographic, social, economic, and housing characteristics.

In other words, while the 2010 Census will show the *number* of people who live in the United States, the American Community Survey shows *how* people live.

Just as development of the ACS required many years, so too will complete adoption of the ACS by researchers, the media, and policy-makers. Although it may take some time to adjust to the changes introduced by the ACS, in time, it should prove to be a highly valuable resource for a wide audience of data users.