

Research & Policy Brief Series

ISSUE NUMBER 24/DECEMBER 2008

Using the American Community Survey (ACS) for Rural or Small Area Research & Policy.

By **Richard Rathge**, **Karen Olson**, and **Ramona Danielson**, North Dakota State University

What is the Issue?

While the decennial U.S. Census is referred to by many as the “gold standard”, data that are collected only once per decade soon become outdated. However, conducting a more frequent national census is cost-prohibitive. The American Community Survey (ACS) was developed to respond to the need for affordable, up-to-date data on U.S. communities. However, while it accomplishes these goals it also presents challenges, particularly with regard to coverage of rural and small areas across the U.S.

What is the ACS?

Historically, the Census Bureau has used a decennial census as its primary means of data collection. The census was conducted in two parts: the short form which enumerated the population and the long form which examined the population's socioeconomic structure. The short form was distributed to all housing units in the nation. It is mandated by the Constitution for political redistricting of the states and reapportionment of the congress. In 2000, the short form contained six population questions (e.g., age, sex, and race) and one housing question related to tenure. The long form contained additional questions on a wide range of social and economic issues. It was distributed to a 1-in-6 sample of addresses. These data provided an empirical basis for many government decisions, functions and policies.

While the Census Bureau will conduct a full count of the population in 2010 as mandated by the Constitution, it will no longer collect information on social and economic characteristics as part of the decennial census. The annual ACS is the Census Bureau's new program that replaces the long form of the decennial census, motivated by the nation's increasingly rapid pace of change, and a corresponding demand for more timely data. The ACS is based on a continuous sampling design with roughly 250,000 monthly surveys which are sent to addresses in the U.S. and Puerto Rico for a total of about 3 million addresses over the calendar year. In order to provide the more timely ACS data, the Census Bureau needed to lower costs by reducing sample size. The ACS's overall annual sampling ratio is about 1-in-40 compared to 1-in-6 for the census long form that was conducted through the year 2000.

Period versus Point Estimates

One of the fundamental differences between the ACS and the census long form is the nature of the data they produce. The census captures a snapshot of the characteristics of people and housing units at one point in time (i.e., a point estimate), typically April 1st of the beginning year of each decade. In contrast, since

the ACS is designed for continuous data collection, it pools responses from 12 months of data collection into one period estimate. These period estimates include changes that occur during the previous 12 months.

For example, in the ACS, persons who are surveyed in January 2007 are asked to report their income over the past 12 months. This means that most of the income they report was obtained in 2006. In contrast, a person who is surveyed in December 2007 would report income generated in 2007. These monthly income responses are pooled together to derive the period estimate for all of 2007 and adjustments are made for inflation. This is very different from a point estimate calculated from the census long form, where data are based on respondents' income earned during the previous calendar year. Hence while census responses refer to the same calendar year period, the ACS may pool data spanning parts of two different calendar years.

The issue of ACS period estimates is further complicated because differential time periods are needed to collect the data. Since it is not financially feasible to collect a large enough sample of addresses in a given year to produce reliable estimates for smaller geographic areas, the ACS has a three-tiered data collection process. Geographic areas of at least 65,000 people are large enough to produce reliable period estimates after 12 months of data collection. In contrast, geographic areas of 20,000 to 65,000 people require 36 months of data collection, and areas smaller than 20,000 people require 60 months of continuous data collection to produce reliable estimates.

The first national release of data from the ACS was in 2005. The first three-year period estimate for places between 20,000 and 65,000 people is scheduled for release in December of 2008. These estimates will represent the 2005-2007 period. Similarly, the three-year period estimate for these places released in 2009 will represent 2006-2008. Table 1 provides a list of geographic units delineated by the Census Bureau and their corresponding period estimates. Researchers using counties as their unit of analysis can see that only 761 of the nation's 3,141 counties will have annual one-year estimates. More than half of counties will have three-year period estimates while the remaining 42 percent which have a population of fewer than 20,000 persons will rely on five-year period estimates. Similarly, only 476 of the nation's 25,161 incorporated places (or designated places) will have one-year period estimates. Researchers conducting neighborhood analysis utilizing tracts or block groups will be restricted to five-year period estimates.

Interpreting Period Estimates

An important challenge for researchers who focus on social and economic change in rural and small areas is to interpret

Table 1: Major Types of Geographic Areas for Which One-Year, Three-Year, and Five-Year Data are Available from the American Community Survey

Type of Geographic Area	Number of Geographic Areas		
	One-Year Period (12 Months): Geographic Areas with a Population of 65,000 or More ¹	Three-Year Period (36 Months): Geographic Areas with a Population of 20,000 or More ¹	Five-Year Period (60 Months): All Geographic Areas ¹
United States	1	1	1
States (and D.C.)	51	51	51
Counties	761	1,811	3,141
Places	476	1,983	25,161
American Indian/Alaska Native Areas	15	41	768
Metropolitan/Micropolitan, Statistical Areas	561	905	923
Congressional Districts	436	436	436
School Districts	879	3,290	14,505
Census Tracts	-	-	65,443
Block Groups	-	-	208,790

¹ The numbers are for geographic areas within the U.S. and do not include areas within Puerto Rico. Also, the numbers shown in the table are based on the latest sources available and may change slightly.

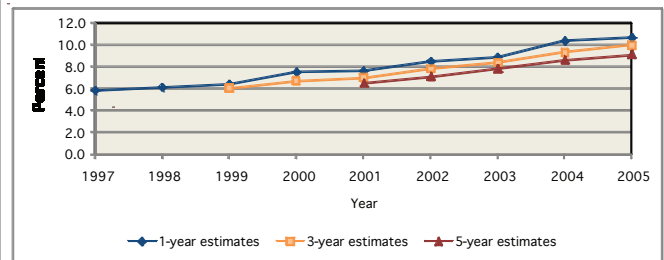
period estimates correctly. They need to remember that these data capture events that occur over the period of the estimate, one, three, or five years, rather than a single point in time (i.e., April 1st) as was previously the case when using data from the census long form. If the characteristic being estimated is very dynamic, the length of the period estimate can be problematic. Two figures using data from the city of Omaha, Nebraska, help illustrate this point¹. In Figure 1, we present the one, three-, and five-year period estimates for the percentage of the population with Hispanic/Latino origin, capturing the systematic gains in each period estimate. A dampening effect appears in the longer period estimates. As a result, using a longer period estimate would likely underestimate the proportion of Hispanic/Latino population.

A greater challenge occurs when the characteristic is both dynamic and fluctuating, such as with poverty rates, as illustrated in Figure 2. Annual estimates show a general decline in the proportion of individuals in poverty from 1997 to 2001, and then an upward swing extending to 2005. The three-year estimate dampens the fluctuation in poverty rates, but still captures the upward swing beginning in 2001. In contrast, the five-year period estimate indicates that the upward swing in poverty commenced in 2003, much later than indicated by the shorter estimates. The differences reflect a significant challenge for data interpretation, especially when the data are used to inform policy decisions and/or to administer public programs.

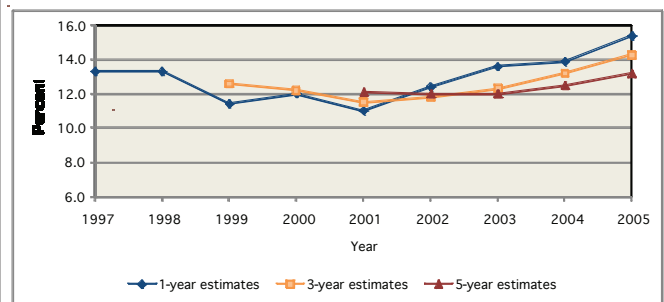
Consequences of Sample Size

Regardless of the length of the period estimate, sample size is another key issue which affects the margin of error (MOE) of ACS estimates. The MOE is the range in which an estimate may actually lie given the confidence level, which, for ACS data, is typically 90 percent. For example, data from the 2007 ACS (i.e., one-year period estimates) for poverty in the Fargo-Moorhead,

¹ The data for the city of Omaha were made available by Dr. Jerry Deichert, Director of Nebraska's State Data Center. We appreciate his willingness to share this information with us.

Figure 1: Percentage of Omaha Population with Hispanic/Latino Origin: 1997 to 2005 ACS

Note: From 1997 to 2005, the ACS surveyed households only and excluded group quarters population.

Figure 2: Percentage of Omaha Population Living in Poverty: 1997 to 2005 ACS

Note: From 1997 to 2005, the ACS surveyed households only and excluded group quarters population.

ND-MN Metropolitan Statistical Area, show an estimated family poverty rate of 8.4 percent, with a 2.3 percent MOE. However, the proportion of families with related children under 5 years of age in poverty (a smaller sample size) is 22.8 percent with a MOE of 11.2 percentage points. This means that there is a 90 percent chance that the true estimate ranges from 11.6 percent to 34.0 percent. Statistics with this magnitude of error will not be useful to legislators attempting to determine whether the prevalence of poverty in their areas is a large and growing problem. Their skepticism of the data will be an important challenge for social scientists using the ACS to overcome.

Implications for Researchers and Policymakers

These issues illustrate just a few of the challenges posed for social scientists planning to use the ACS in policy analysis. The most daunting may be how researchers and policymakers use and interpret the data, not only for themselves but for the broader audiences and constituents they serve. The research community often sets the standards for data use and interpretation, and therefore must determine the conventions they wish to encourage regarding the use of ACS period estimates especially for smaller places where estimates represent multiple years of data collection. The Census Bureau is assisting in this task by providing educational materials (located at http://www.census.gov/acs/www/UseData/Compass/compass_series.html). Users need to appropriately and effectively use ACS data and avoid the skepticism that may arise because of limitations. One possible solution is to encourage triangulation of longer period estimates with other administrative records. This will increase confidence in the accuracy of ACS estimates and/or it will alert users when ACS estimates are unreliable. ▲

