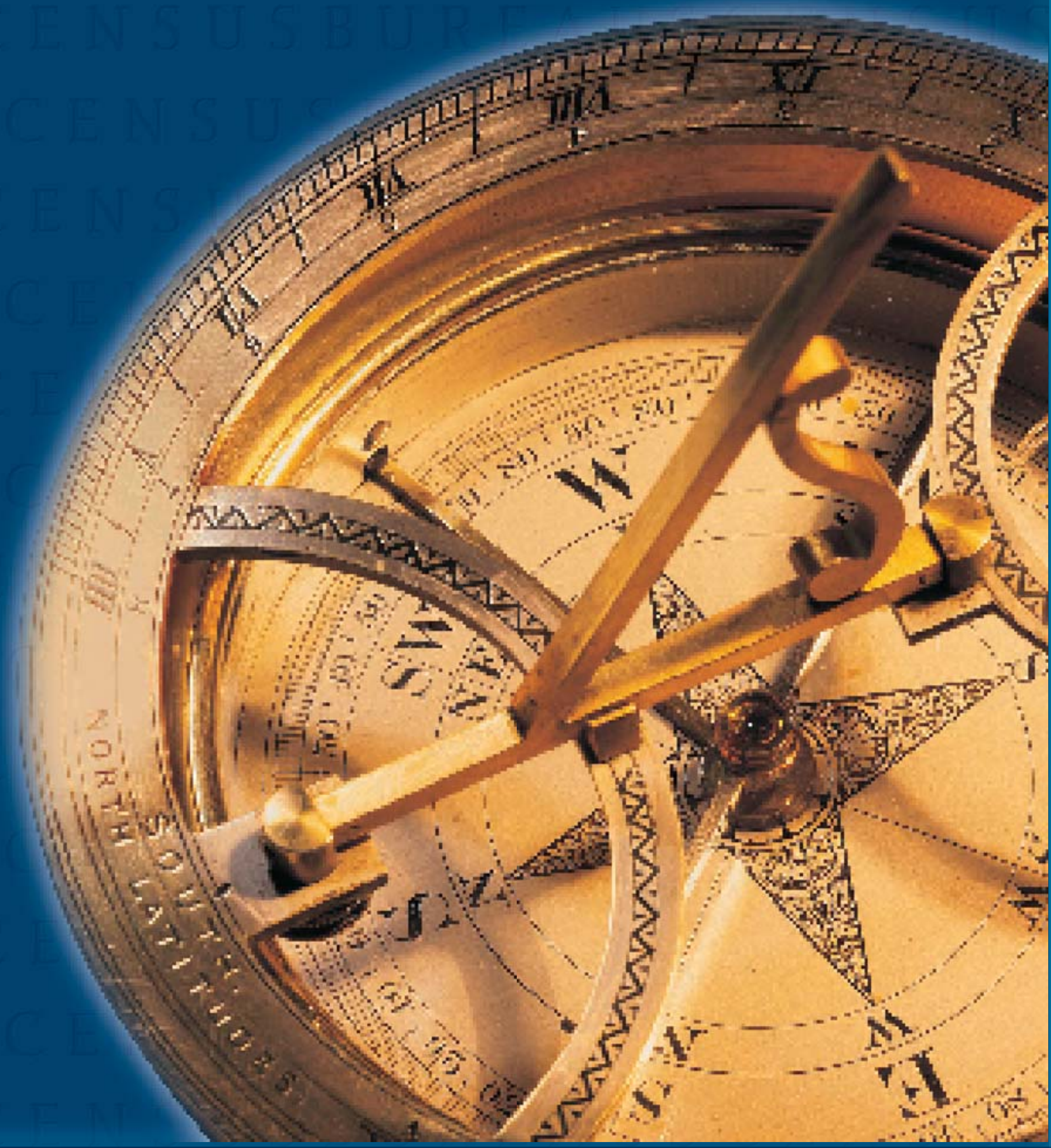
A Compass for Understanding and Using American Community Survey Data

What the Business Community Needs to Know

Issued
October 2008



USCENSUSBUREAU

Helping You Make Informed Decisions

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

United States Census 2010

Acknowledgments

Ken Hodges, Nielsen Claritas, drafted this handbook for the U.S. Census Bureau's American Community Survey Office. **Kennon R. Copeland** and **John H. Thompson** of National Opinion Research Center at the University of Chicago drafted the technical appendixes. **Edward J. Spar**, Executive Director, Council of Professional Associations on Federal Statistics, **Frederick J. Cavanaugh**, Executive Business Director, Sabre Systems, Inc., **Susan P. Love**, Consultant, **Linda A. Jacobsen**, Vice President, Domestic Programs, Population Reference Bureau, and **Mark Mather,** Associate Vice President, Domestic Programs, Population Reference Bureau, provided initial review of this handbook.

Deborah H. Griffin, Special Assistant to the Chief of the American Community Survey Office, provided the concept and directed the development and release of a series of handbooks entitled *A Compass for Understanding and Using American Community Survey Data.* **Cheryl V. Chambers**, **Colleen Flannery**, **Cynthia Davis Hollingsworth**, **Susan L. Hostetter**, **Pamela D. Klein**, **Clive R. Richmond, Enid Santana**, **Anna M. Owens**, and **Nancy K. Torrieri** contributed to the planning and review of this handbook series.

The American Community Survey program is under the direction of **Arnold A. Jackson**, Associate Director for Decennial Census, **Daniel H. Weinberg**, Assistant Director for the American Community Survey and Decennial Census, and **Susan Schechter**, Chief, American Community Survey Office.

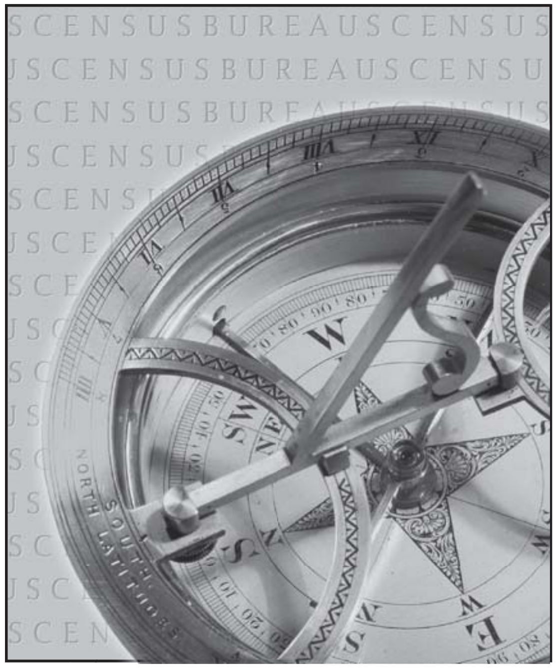
Other individuals who contributed to the review and release of these handbooks include **Dee Alexander**, **Herman Alvarado**, **Mark Asiala**, **Frank Ambrose**, **Maryam Asi**, **Arthur Bakis**, **Genora Barber**, **Michael Beaghen**, **Judy Belton**, **Lisa Blumerman**, **Scott Boggess**, **Ellen Jean Bradley**, **Stephen Buckner**, **Whittona Burrell**, **Edward Castro**, **Cheryl Chambers**, **Gary Chappell**, **Michael Cook**, **Russ Davis**, **Carrie Dennis**, **Jason Devine**, **Joanne Dickinson**, **Barbara Downs**, **Maurice Eleby**, **Colleen Flannery**, **Sirius Fuller**, **Dale Garrett**, **Yvonne Gist**, **Marjorie Hanson**, **Greg Harper**, **William Hazard**, **Steve Hefter**, **Douglas Hillmer**, **Frank Hobbs**, **Cynthia Davis Hollingsworth**, **Susan Hostetter**, **Todd Hughes**, **Trina Jenkins**, **Nicholas Jones**, **Anika Juhn**, **Donald Keathley**, **Wayne Kei**, **Karen King**, **Debra Klein**, **Pamela Klein**, **Vince Kountz**, **Ashley Landreth**, **Steve Laue**, **Van Lawrence**, **Michelle Lowe**, **Maria Malagon**, **Hector Maldonado**, **Ken Meyer**, **Louisa Miller**, **Stanley Moore**, **Alfredo Navarro**, **Timothy Olson**, **Anna Owens**, **Dorothy Paugh**, **Marie Pees, Marc Perry**, **Greg Pewett**, **Roberto Ramirez**, **Dameka Reese**, **Katherine Reeves**, **Patrick Rottas**, **Lil Paul Reyes**, **Clive Richmond**, **Merarys Rios**, **J. Gregory Robinson**, **Anne Ross**, **Marilyn Sanders**, **Enid Santan-Ortiz**, **Nicole Scanniello**, **David Sheppard**, **Joanna Stancil**, **Michael Starsinic**, **Lynette Swopes**, **Anthony Tersine**, **Nancy Torrieri**, **Carrie Werner**, **Edward Welniak**, **Andre Williams**, **Steven Wilson**, **Kai Wu**, and **Matthew Zimolzak**.

Linda Chen and **Amanda Perry** of the Administrative and Customer Services Division, **Francis Grailand Hall**, Chief, provided publications management, graphics design and composition, and editorial review for the print and electronic media. **Claudette E. Bennett**, Assistant Division Chief, and **Wanda Cevis**, Chief, Publications Services Branch, provided general direction and production management.

A Compass for Understanding and Using American Community Survey Data

Issued October 2008

What the Business Community Needs to Know



U.S. Department of Commerce
Carlos M. Gutierrez,
Secretary

John J. Sullivan,
Deputy Secretary

Economics and Statistics Administration
Cynthia A. Glassman,
Under Secretary for Economic Affairs

U.S. CENSUS BUREAU
Steve H. Murdock,
Director

Suggested Citation

U.S. Census Bureau,
A Compass for Understanding and Using American Community Survey Data: What the Business Community Needs to Know
U.S. Government Printing Office,
Washington, DC,
2008.

ECONOMICS
AND STATISTICS
ADMINISTRATION

Contents

Foreword

The American Community Survey (ACS) is a nationwide survey designed to provide communities with reliable and timely demographic, social, economic, and housing data every year. The U.S. Census Bureau will release data from the ACS in the form of both single-year and multiyear estimates. These estimates represent concepts that are fundamentally different from those associated with sample data from the decennial census long form. In recognition of the need to provide guidance on these new concepts and the challenges they bring to users of ACS data, the Census Bureau has developed a set of educational handbooks as part of *The ACS Compass Products*.

We recognize that users of ACS data have varied backgrounds, educations, and experiences. They need different kinds of explanations and guidance to understand ACS data products. To address this diversity, the Census Bureau worked closely with a group of experts to develop a series of handbooks, each of which is designed to instruct and provide guidance to a particular audience. The audiences that we chose are not expected to cover every type of data user, but they cover major stakeholder groups familiar to the Census Bureau.

- General data users
- High school teachers
- Business community
- Researchers
- Federal agencies
- Media
- Congress
- Puerto Rico Community Survey data users (in Spanish)
- Public Use Microdata Sample (PUMS) data users
- Users of data for rural areas
- State and local governments
- Users of data for American Indians and Alaska Natives

The handbooks differ intentionally from each other in language and style. Some information, including a set of technical appendixes, is common to all of them. However, there are notable differences from one handbook to the next in the style of the presentation, as well as in some of the topics that are included. We hope that these differences allow each handbook to speak more directly to its target audience. The Census Bureau developed additional *ACS Compass Products* materials to complement these handbooks. These materials, like the handbooks, are posted on the Census Bureau's ACS Web site: <www.census.gov/acs/www>.

These handbooks are not expected to cover all aspects of the ACS or to provide direction on every issue. They do represent a starting point for an educational process in which we hope you will participate. We encourage you to review these handbooks and to suggest ways that they can be improved. The Census Bureau is committed to updating these handbooks to address emerging user interests as well as concerns and questions that will arise.

A compass can be an important tool for finding one's way. We hope *The ACS Compass Products* give direction and guidance to you in using ACS data and that you, in turn, will serve as a scout or pathfinder in leading others to share what you have learned.

What Is the American Community Survey?

The American Community Survey (ACS) is the biggest change in the U.S. decennial census since businesses became major users of decennial census data in the early 1970s. In recent decades the census has consisted of a “short form” that was sent to most households, asking just a few basic questions, and a “long form,” sent to a sample of households, gathering data on many detailed characteristics. Long-form data have been especially popular in business applications as the source of information on income, education, language, commuting, housing value, and many other items. The provision of data for very small areas has been a key to much of this popularity.

Beginning with the 2010 Census, the census will collect information only on the basic short-form items. There will be no long form in 2010, but the data that businesses have come to expect will still be provided. Instead of collecting the data as part of the 2010 Census, the Census Bureau now is collecting this information throughout the decade using the ACS. “Long-form replacement” is a key ACS objective, but ACS data are not a literal replacement for long-form data. Content is similar to the long form, but ACS data are different enough that users should not simply plug them into existing applications without a basic understanding of this new data source. The ACS provides a new, wide range of data products. To make effective use of the ACS, businesses need to understand how ACS data are collected and reported, and their advantages and limitations relative to the long-form data previously collected during the decennial census. A glossary and a series of technical appendixes—for those interested in more advanced ACS issues—are included at the back of this handbook.

Note to new data users:

This handbook makes frequent references to census long-form sample data, and it describes how ACS data differ from data previously provided by the census. These comparisons are for the benefit of those who have been using census sample data and who need to transition to the ACS. In contrast, new users can start fresh, focusing on the ACS for what it is—a very large survey that will provide regularly updated estimates of a wide variety of demographic and socioeconomic variables that are valuable for many business applications.

While new users need not be concerned with comparisons to traditional long-form data, they will want to pay close attention to the handbook’s explanation of the nature of ACS estimates. Of particular importance are the descriptions of ACS “period estimates,” which are unlike the “point-in-time” estimates provided by most data sources, and the measures of statistical uncertainty involved in ACS estimates.

Why Is the American Community Survey Important?

The ACS is important to business for the same reasons that census long-form data have been important. The census collects information for federal program purposes, but businesses have been creative in finding prolific uses for census data—especially data from the long form. Census data have become a foundation for the research, planning, and evaluation work of many businesses.

Business uses of demographic and socioeconomic data fall into two major categories—market/site evaluation and consumer segmentation. These applications often require data for very small areas, and the census has been the most authoritative source of small area data available. And while many businesses now use sophisticated and proprietary site and segmentation products provided by private firms, these products are built largely on a statistical and geographic foundation provided by the census. The Decennial Census Program, now including the American Community Survey, remains a critical resource to the companies that build site and segmentation tools and to the companies that use them.

Examples of types of questions that can be addressed by the ACS:

- How many high-income households with children live near a specific store location?
- Where can one find concentrations of people of Polish ancestry in the Pittsburgh metropolitan area?
- How has the age composition of Phoenix changed in recent years?
- Which Midwestern metropolitan areas have the highest percent of foreign-born population?
- Where can one find concentrations of housing held for seasonal, recreational, or occasional use?
- Which Chicago neighborhoods have concentrations of people who speak Spanish at home?
- How many workers in Atlanta leave home for work before 6 a.m.?

The Bottom Line on ACS Advantages and Limitations

One could fill a book on the topic of ACS advantages and limitations relative to the decennial census long form. ACS data involve some basic tradeoffs that can be summarized as follows.

ACS data have the advantage of being more current than census long-form data because they are collected and disseminated throughout the decade. Unlike the census, ACS data will never be 10 years out of date. With content similar to the long form, the ACS will provide updates of items, such as household composition (for example, single-parent households with children), that are often used by businesses but have not been frequently updated for small areas. There is also an advantage in quality, as information from nonresponding ACS households is gathered by highly trained and experienced interviewers, in contrast with the temporary workers who followed up on households that did not return census long forms.

Several factors offset these ACS advantages. First, the ACS sample is smaller than that of the Census 2000 long form. About 18.3 million housing units were in the Census 2000 long-form sample, while 5 years of ACS sample addresses will sum to around 15 million. Because the ACS follows up on only a subsample of nonresponding addresses, the number of housing units actually interviewed over 5 years is expected to be about 10 million. Second, the once-per-decade long-form data were controlled to actual census counts of population and housing, while the annual ACS data are weighted to conform to the Census Bureau population and housing estimates that update the census counts every year. The uncertainty in the Census Bureau's population and housing estimates contributes to uncertainty in ACS data. (Appendix 7 describes the implications of the use of controls based on Census Bureau estimates.) Third, while long-form data described characteristics approximating a point in time or a fixed reference period, ACS data reflect population and housing characteristics over a period of time—1 year, 3 years, or 5 years. Consequently, the interpretation of ACS estimates is not as straightforward as some users might like.

The effects of these differences may be largest for the small area data that are common to business applications. Individual users will have to weigh these impacts to determine the suitability of ACS data for their applications. For example, the impact on site evaluation work may be minimal, as unreliable block group estimates are aggregated to trade areas large enough to have low sampling error. But segmentation applications that depend on accurate statistical portraits for specific small areas—such as individual block groups—may be impaired.

Where characteristics such as age, income, length of residence, or ethnicity are involved, private sector products built from consumer household databases might be preferred as a means to target individual block groups or individual households. However, for characteristics such as language, employment status, disability, and many others, the ACS (or products built from the ACS) may be the only source of small area data. Users of these data for individual small areas need to understand the relatively large sampling errors that can be involved and that are reported with all ACS estimates.

In assessing the suitability of small area ACS data for their applications, business users should keep in mind that long-form data for individual tracts and block groups also were subject to relatively large sampling errors and were best used for aggregated areas. Long-form data errors were less apparent, as margins of error were not reported with the published estimates, and the once-per-decade schedule masked the fluctuations that would have been observed if long-form data had been provided annually. Knowing the variability and reliability of the data, although it may require more understanding and analysis, should help businesses arrive at better decisions.

ACS Data Collection

The Census Bureau started testing the ACS in 1996 and went to the full nationwide sample in 2005. In contrast to the once-per-decade census, the ACS collects data continuously during the calendar year, with about 250,000 housing units added to the ACS sample each month.

The ACS questionnaire is similar to that of the census long form, but some questions are asked differently, and content has evolved in recent years. Copies of ACS questionnaires for recent years can be reviewed at the ACS Web site. Go to <www.census.gov/acs/www/SBasics/SQuest/SQuest1.htm> to find the link.

The ACS is a mailout/mailback survey. Each month the ACS questionnaire is sent to about 250,000 addresses sampled from the Census Bureau's Master Address File. About 6 weeks later, the Census Bureau follows up by telephone, when possible, on addresses that have not returned a completed questionnaire. For addresses that still have not responded, a subsample, ranging between 1-in-2 and 1-in-3, is selected for additional follow-up. Sample addresses that could not be mailed are subsampled at a rate of 2-in-3. Census Bureau field interviewing staff visit these addresses in person and conduct interviews.

Starting in 2006, in a separate operation, data are collected from people who live in group quarters facilities, such as college/university student housing, military barracks, nursing home facilities, and correctional facilities. In this process, a sample of group quarters facilities is selected. Census Bureau field staff visit each selected facility, interview a contact person (such as an administrator) to collect facility level information, and then select and interview a sample of residents living at the facility. The ACS group quarters questionnaires can be reviewed at <www.census.gov/acs/www/SBasics/SQuest/SQuest1.htm>.

Response to the ACS is required by law, just as it is for the decennial census. As with the census, responses to the ACS are strictly confidential, and the Census Bureau takes great care to preserve this confidentiality in ACS data products.

How ACS Data Are Provided

Although ACS data collection is going on nearly every day, specific data products are updated once each year and released over several months, beginning in late August through December. Three basic types of data will be released—estimates reflecting 1 year of data collection, 3 years of data collection, and 5 years of data collection. One year of data collection reflects an initial sample of around 3 million addresses, 3 years nearly 9 million, and the 5-year data reflect close to 15 million addresses. However, because the ACS follows up on only a subsample of nonresponding addresses, the number of completed interviews is smaller than the initial sample sizes indicated above, close to 2 million housing unit interviews and 150,000 interviews with group quarters residents every year.

The population size of a geographic area determines the type of ACS estimates that are provided. Estimates that are based on 1 year of data collection are provided for geographic areas with populations of 65,000 or more and 3-year estimates are provided for areas of 20,000 or more. Only 5-year ACS estimates will be provided for areas of the smallest population size—including small towns, census tracts, and block groups. It is these 5-year data that come closest to replacing the small area census long-form sample data that many businesses have used in recent decades.

ACS data are provided (or will be provided) for a variety of geographic areas including those identified in Table 1 below. This table gives you a sense of which types of areas are large enough to receive 1-year and 3-year estimates. As the number of areas in column 1 makes clear, the vast majority of geographic areas will receive only 5-year estimates.

The Census Bureau plans to release new or updated, sets of 1-year, 3-year, and 5-year ACS data every year. Since 2005 was the first year of ACS data collection at full sample, the first set of 1-year estimates was released in 2006. The first set of 3-year estimates will be released in 2008, and the first set of 5-year estimates will be released in 2010. Following 2010, the Census Bureau will release new 1-year, 3-year, and 5-year estimates every year.

Table 2 summarizes the data collection periods and year of release for ACS data products (through 2011), as data collection builds to permit the annual release of 1-year, 3-year, and 5-year data.

Table 1. **Major Geographic Areas and Type of ACS Estimates Received**

Type of geographic area	Total number of areas	Percent of total areas receiving . . . 1-year, 3-year, & 5-year estimates	3-year & 5-year estimates only	5-year estimates only
States and District of Columbia	51	100.0	0.0	0.0
Congressional districts	435	100.0	0.0	0.0
Public Use Microdata Areas*	2,071	99.9	0.1	0.0
Metropolitan statistical areas	363	99.4	0.6	0.0
Micropolitan statistical areas	576	24.3	71.2	4.5
Counties and county equivalents	3,141	25.0	32.8	42.2
Urban areas	3,607	10.4	12.9	76.7
School districts (elementary, secondary, and unified)	14,120	6.6	17.0	76.4
American Indian areas, Alaska Native areas, and Hawaiian homelands	607	2.5	3.5	94.1
Places (cities, towns, and census designated places)	25,081	2.0	6.2	91.8
Townships and villages (minor civil divisions)	21,171	0.9	3.8	95.3
ZIP Code tabulation areas	32,154	0.0	0.0	100.0
Census tracts	65,442	0.0	0.0	100.0
Census block groups	208,801	0.0	0.0	100.0

* When originally designed, each PUMA contained a population of about 100,000. Over time, some of these PUMAs have gained or lost population. However, due to the population displacement in the greater New Orleans areas caused by Hurricane Katrina in 2005, Louisiana PUMAs 1801, 1802, and 1805 no longer meet the 65,000-population threshold for 1-year estimates. With reference to Public Use Microdata Sample (PUMS) data, records for these PUMAs were combined to ensure ACS PUMS data for Louisiana remain complete and additive.

Source: U.S. Census Bureau, 2008. This tabulation is restricted to geographic areas in the United States. It was based on the population sizes of geographic areas from the July 1, 2007, Census Bureau Population Estimates and geographic boundaries as of January 1, 2007. Because of the potential for changes in population size and geographic boundaries, the actual number of areas receiving 1-year, 3-year, and 5-year estimates may differ from the numbers in this table.

Table 2. **Release Schedule for ACS Data**

Data product	Population threshold	Year of Data Release 2006	2007	2008	2009	2010	2011	2012	2013
		Year(s) of Data Collection							
1-year estimates	65,000+	2005	2006	2007	2008	2009	2010	2011	2012
3-year estimates	20,000+			2005–2007	2006–2008	2007–2009	2008–2010	2009–2011	2010–2012
5-year estimates	All areas*					2005–2009	2006–2010	2007–2011	2008–2012

*Five-year estimates will be available for areas as small as census tracts and block groups.
Source: U.S. Census Bureau.

What about ZIP Codes?

The Census Bureau also plans to release ACS data for ZIP Code Tabulation Areas (ZCTAs). First defined by the Census Bureau for Census 2000 data products, ZCTAs are aggregations of census blocks that approximate areas corresponding to ZIP Codes. The emphasis is on "approximate" because ZIP Codes are numeric codes for mail delivery and often do not have clear geographic boundaries. Because most ZCTAs are small, the Census Bureau will only release them in the form of 5-year estimates. Because ZIP Codes are subject to change, the Census Bureau plans to update ZCTA definitions periodically. However, ZCTA definitions will not be updated every year, so users will need to pay close attention to the "vintage" of the ZCTA definitions in ACS products and understand that these definitions may differ from ZIP Codes as defined by the U.S. Postal Service or by private data suppliers.

How the ACS Differs From Other Data Sources

While "long-form replacement" is an ACS objective, it is not a replacement in a literal sense, as there are fundamental differences between ACS and census long-form data. The census long-form sample was a large survey taken once a decade as part of the decennial census, while the ACS is a large survey conducted throughout the decade—with a new sample of around 250,000 addresses introduced each month. The collected information is aggregated and processed once a year, and the resulting ACS estimates reflect population and characteristics over periods of time—either 1, 3, or 5 years, as described above. The complications of these "period estimates" are described in a later section.

The ACS uses a different definition than the census of who lives in a housing unit. The census counts people according to their "usual residence," or where they live most of the year. The ACS includes people according to a "2-month rule" of "current residence." People are considered residents of a sample unit if their length of stay has been or will be more than 2 months when contacted by the ACS. This difference is important for businesses that need to know if their data include seasonal populations who reside in an area only a few months of the year. For the census, the answer is "No," but for the ACS, the answer is not a clear "Yes" or "No." For example, in parts of Florida, ACS data collection in winter months will interview people who reside there only during the winter if their stay exceeds 2 months. ACS data products describe the characteristics of the population of an area throughout the entire year (or years) of data collection based on stays of more than 2 months, and thus reflect a mix of seasonal highs and lows.

The census is focused on population counts, and for purposes of congressional apportionment and redistricting, it is important that people be counted only once and in the right place. *The ACS is not about population counts but about characteristics.* For example, the ACS does not determine the estimate of the total population of Fargo, North Dakota. The total population for Fargo provided in ACS tables is based on the Census Bureau's Population Estimates Program. However, if one wants to know the percent of Fargo's population that is widowed or divorced, the ACS is the source for that estimate.

Most ACS characteristics are similar in definition to their counterparts from the census long-form sample, but users should be aware of some differences. Because so many businesses rely on estimates of income and home value, these users need to understand how the ACS handling of these estimates differs from what they have been using from the long form. The Census 2000 long form asked respondents to report income received during the previous calendar year, so income reported by Census 2000 reflected income received during 1999. In contrast, the ACS asks respondents to report income received during the previous 12 months. With data collected throughout the year, the incomes reported reflect a range of 12 different 12-month income reference periods. People interviewed in March 2007 reported income received from March 2006 through February 2007, while those interviewed in November 2007 reported income received from November 2006 through October 2007. Among "2007" ACS respondents, reference periods span 23 months (from January 2006 through November 2007).

When dealing with ACS income data, it is important to keep in mind that the dollar amounts are inflation adjusted to reflect changes in the national consumer price index (CPI)—with results expressed in dollar values of the most current data collection year. For example, each of the 12 different income reference periods for 2007 ACS respondents are inflation adjusted to represent 2007 dollars. When combining multiple ACS samples, as is done to create 3-year and 5-year estimates, the earlier year ACS respondent incomes are further inflation adjusted to the most recent ACS year of the multiyear period. Five-year income estimates based on data collected from 2005 through 2009 will

Table 3. **Differences in Income Reported by Census 2000 and the ACS**

	Census 2000 Long Form	**ACS 2005–2009**
Total Households	As of April 2000	Averaged over collection period
Collection period	Spring and summer, 2000	Every month—January 2005 through December 2009
Income reporting	Income received in 1999	Income received during 12-month period prior to interview
Inflation adjustments	None	Income inflation adjusted to the most current data collection year

Source: U.S. Census Bureau.

reflect the incomes reported during that period adjusted to 2009-dollar values. Table 3 summarizes the differences between income data from Census 2000 and 5-year ACS estimates reflecting data collected during 2005 through 2009. Similar adjustments are made for 3-year estimates.

Note that income is the only variable for which such adjustments are made for the 1-year estimates. For all other data items reported as dollar values, inflation adjustments are made only for the 3-year and 5-year estimates. The adjustments are made to reflect dollar values of the latest collection year in the period.

See Appendix 2 for additional detail on differences between ACS and decennial census sample data. Appendix 5 provides detail on the use of dollar-denominated ACS data.

The ACS also differs in important ways from the Current Population Survey (CPS), which also is used as a source of demographic and socioeconomic information. The ACS sample of nearly 3 million addresses per year is much larger than the CPS sample of about 50,000 households per month. While both the CPS and ACS collect data throughout the year, the CPS collects demographic and income detail in the Annual Social and Economic Supplement (ASEC), based on a sample of about 100,000 addresses initially interviewed in CPS during February, March, and April. With its larger sample, the ACS is designed to provide data for much smaller areas than the CPS. However, for some topics, the CPS asks more detailed questions and may produce results that are preferable for some applications. Thus for seemingly similar data items, such as median income or poverty rates, the ACS and the CPS will sometimes report different results at national and state levels.

What ACS Data Products Are Available?

Despite the differences between ACS and census long-form data, experienced census data users should feel very much at home with ACS data products and new users should find them easy to get to know. All ACS products can be accessed through the American FactFinder on the Census Bureau's Web site: <http://factfinder.census.gov>.

Current products include **Detailed Tables** that provide the most detailed information on all ACS subjects for all geographic levels. Users can think of these tables as similar to those from the Census 2000 Summary Files (such as SF1 and SF3). One difference is that the ACS tables report measures of sampling error called *margins of error* for every estimate, while the Census 2000 sample data products provided no such measures. The exception is ACS estimates that are controlled to independent Census Bureau estimates. Measures of sampling error are not provided because these estimates are not subject to sampling error.

The detailed table below shows the estimates and margins of error of sex by educational attainment for Los Angeles County, California, in 2006. The data are produced for the population aged 25 and over.

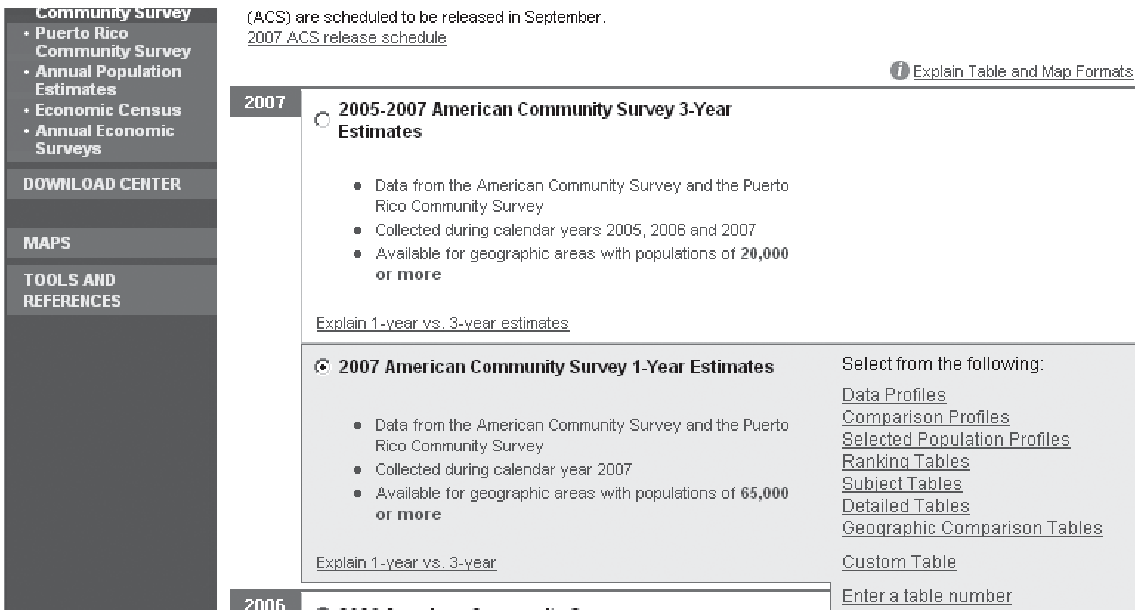
Figure 1. **Example of a Detailed Table**

	Los Angeles County, California	
	Estimate	Margin of Error
Total:	6,243,935	*****
Male:	3,032,707	*****
No schooling completed	63,407	+/-4,984
Nursery to 4th grade	73,569	+/-4,809
5th and 6th grade	184,779	+/-7,598
7th and 8th grade	93,525	+/-5,412
9th grade	94,346	+/-5,378
10th grade	56,807	+/-4,609
11th grade	75,834	+/-4,273
12th grade, no diploma	108,697	+/-6,121
High school graduate (includes equivalency)	701,506	+/-12,009
Some college, less than 1 year	135,157	+/-6,023
Some college, 1 or more years, no degree	393,413	+/-8,771
Associate's degree	182,550	+/-7,448
Bachelor's degree	555,224	+/-10,556
Master's degree	185,687	+/-6,764
Professional school degree	82,565	+/-4,360
Doctorate degree	45,641	+/-3,362
Female:	3,211,228	*****
No schooling completed	85,803	+/-4,933
Nursery to 4th grade	84,756	+/-4,730
5th and 6th grade	204,150	+/-6,973
7th and 8th grade	103,825	+/-5,779
9th grade	89,975	+/-4,761
10th grade	64,949	+/-4,274
11th grade	80,289	+/-5,536
12th grade, no diploma	95,153	+/-4,231
High school graduate (includes equivalency)	737,600	+/-13,771
Some college, less than 1 year	169,101	+/-6,852
Some college, 1 or more years, no degree	400,722	+/-9,780
Associate's degree	236,045	+/-7,489
Bachelor's degree	587,710	+/-11,190
Master's degree	189,826	+/-7,329
Professional school degree	57,425	+/-3,912
Doctorate degree	23,899	+/-2,475

Source: U.S. Census Bureau, 2006 American Community Survey, accessed at <http://factfinder.census.gov>.

Figure 2 is a screenshot of the American FactFinder data sets page. This page lists the specific ACS data sets (for example, 2007 ACS or 2005–2007 ACS) and the associated ACS data products. Some of these products are described in greater detail below.

Figure 2. **ACS Data Sets Page**



Source: U.S. Census Bureau, American FactFinder, accessed at <http://factfinder.census.gov>.

- **Data profiles** are preformatted reports summarizing characteristics for selected areas. They are similar to the data profiles provided for Census 2000 data. For all areas meeting the required population thresholds, one can find data profile reports summarizing information on the following topics, as well as others. **Narrative profiles** can be accessed from these data profiles and provide information in a user-friendly text and graphic format that puts various topics found in the data profile into words for the general user.
    - o Social: education, marital status, household type, grandparents
    - o Economic: income, employment, occupation, commuting
    - o Housing: housing value, owner/renter, mortgage status
    - o Demographic: age, sex, race, Hispanic origin
- **Selected population profiles** summarize basic data for a generous list of race, Hispanic origin, and ancestry groups. For example, a user could quickly access a report summarizing the characteristics of people who marked both White and Asian in responding to the race question. Beginning in 2008, these selected population profiles will include country of birth.
- **Comparison profiles** make it easy to compare data for 2 sequential years. They are new in 2007 and will present 2006 ACS and 2007 ACS estimates and display the results of statistical testing to assess if the differences in these 2 years are statistically different.
- **Ranking tables** allow users to rank states (plus the District of Columbia and Puerto Rico) on key characteristics. For example, a financial services company that organizes sales territories by groups of states might rank states on variables such as percent of households with retirement income or the median value of owner-occupied housing units.
- **Subject tables** provide expanded detail in tables arranged by subject. For example, a business targeting a Spanish-speaking clientele could define the area of interest and then choose among a variety of tables presenting data related to language spoken at home.
- **Detailed tables** provide the most detailed information on all ACS subjects for all geographic levels. Users can think of these tables as similar to those from the Census 2000 Summary Files (such as SF1 and SF3).

- **Geographic comparison tables** complement the ranking tables by providing data with more geographic detail. For example, if the business described above is surprised by Florida's ranking on median home value, it can run a geographic comparison table to see how median value is distributed across the state's counties.

- **Thematic maps** present the state level information from ranking tables on maps.

- The **ACS Summary File (SF)** includes all ACS detailed tables for all published geographic levels. Businesses that use large quantities of ACS data, or that use ACS data in the development of nationwide marketing information products, may find the Summary File the most efficient way to acquire and store the ACS data they use. Unlike decennial census products, which provide multiple summary files (such as Summary Files 1, 2, 3, and 4), there is only one ACS Summary File.

- **Public Use Microdata Samples (PUMS)** provide anonymous respondent or record level data for a sample of all housing units and residents of group quarters interviewed by the ACS. To protect confidentiality, PUMS data are not provided for small areas—only for large areas known as Public Use Microdata Areas (PUMAs). PUMAs have populations of about 100,000. Because individual level responses are provided, PUMS enables users to generate tables that are not published by the Census Bureau. For example, a research firm developing weights for use in a survey of Asian householders might need estimates of Asian householders by age and place of birth for all states. The ACS detailed tables do not provide this cross-tabulation, but the research firm can produce it from the ACS PUMS product. PUMS also provides the option of producing separate tables for people living in households and people living in group quarters.

  As with traditional census products, ACS PUMS data involve a tradeoff. PUMS data provide great flexibility in data content—enabling users to build data tables that are not available in the published data products (such as the detailed tables). However, the tradeoff is that PUMS data provide much less geographic detail than the published products.

- **Custom tabulations** are an option for users needing data not provided by ACS detailed tables, and which cannot be produced with PUMS. For example, a publisher seeking to identify concentrations of highly educated foreign-born women with children might be disappointed to find that the ACS does not provide a detailed table with this combination of characteristics. If the publisher needs the data for counties or census tracts (to aggregate to market areas), the data cannot be produced from PUMS. A user in this situation can pay the Census Bureau to produce a "custom tabulation" of the data from the ACS detail files. Although attractive for some users, this option involves time and expense. The requested tabulation must be approved by the Census Bureau's Disclosure Review Board to ensure that it does not put the confidentiality of individual data at risk. To further guard against disclosure, custom tabulation data are subject to rounding and suppression.

Different ACS products will appeal to different businesses. Many ACS products function best as a lookup tool. The user has a few areas in mind, wants specific data and maps for those areas, and runs the relevant ACS product. For example, a small Des Moines, Iowa, based retailer could go to the Census Bureau Web site (see details below on how to access the Web site) and obtain selected detailed tables, data profiles, and subject tables for the Sioux City, Omaha, and Lincoln metropolitan areas to determine which is the best candidate for an initial expansion. However, companies that need continuous and complex analyses of markets across the United States would need to acquire large volumes of ACS data—either from a download center provided through American FactFinder or the ACS Summary File. A real estate firm could download ACS data every year to help monitor the demographics in a large number of sites defined as aggregations of block groups. A financial services provider with data on the age and mortgage status of its customers might use ACS PUMS data to build a model suggesting the likelihood that its customers have children approaching college age. Another firm might use ACS data on housing costs as a percent of income to identify neighborhoods where homeowners face difficulty in making house payments.

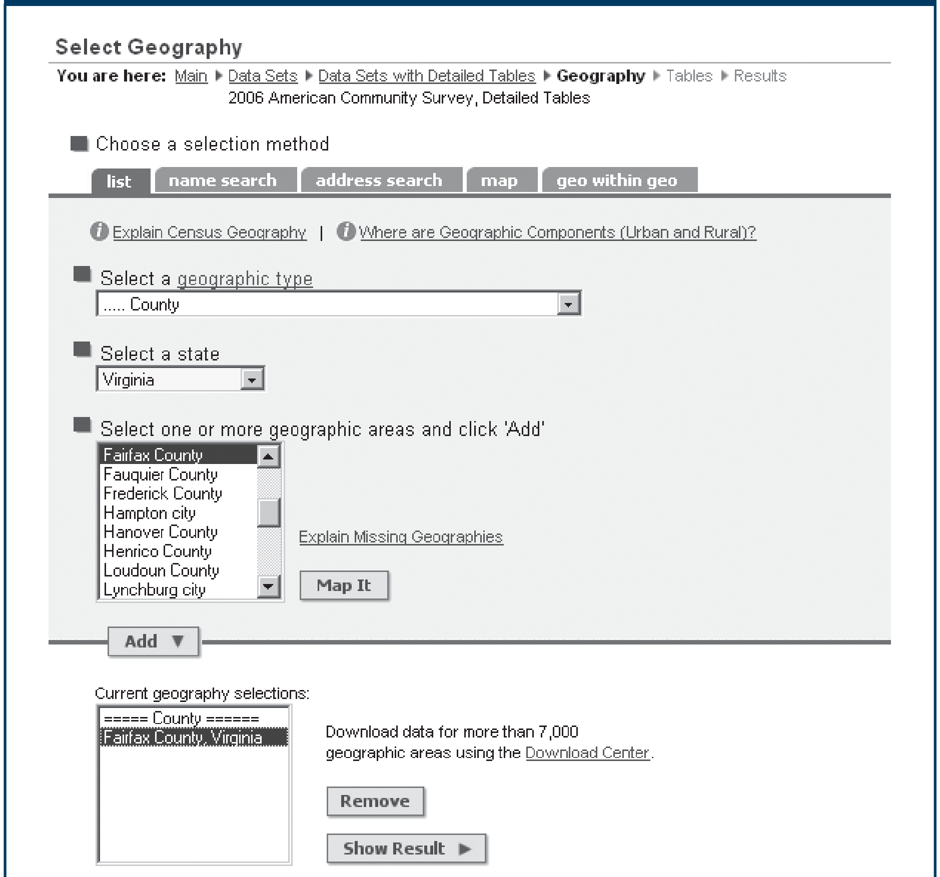
How to Access ACS Data

It remains to be seen how private sector data firms will provide and incorporate ACS data into their products, but data already are available on the Census Bureau's Web site <www.census.gov>. On the Census Bureau homepage, clicking "American Community Survey" takes the data user to the ACS "main page," which provides links to a wide range of information, from introductory material to product descriptions and release schedules to in-depth explanations of topics including data collection and error measures.

Once familiar with the ACS and its products, many users will access ACS data through the American FactFinder, the Census Bureau Web site's function for providing access to data products. The American FactFinder main page can be accessed from the Census Bureau Web site's homepage or at <http://factfinder.census.gov>.

Clicking on "Data Sets" and then on "American Community Survey" opens a page where users choose the desired ACS data set and specific data product (such as detailed tables, ranking tables, subject tables, thematic maps). Figure 2 displays the 2007 Data Sets page. Note that the user by default is directed to the most recent data set. In this example, the 2005–2007 ACS 3-year estimates are shown as the default. This is what users will see when these 3-year estimates are released in December 2008. Using the radio buttons, users can change this selection to another data set including the 2007 ACS 1-year estimates or 1-year estimates for 2005 or 2006. Depending on the product selected, the user is then prompted for the specific geographic area of interest and the specific table or a keyword or subject in order to search for a table. Figure 3 is a screenshot of the page within American FactFinder that prompts the user to identify the geographic area.

Figure 3. **Selecting Geographies in American FactFinder**



Source: U.S. Census Bureau, American FactFinder, accessed at <http://factfinder.census.gov>.

See below for a step-by-step example of how to access an ACS table from the American FactFinder.

To retrieve an ACS Detailed Table, follow the steps below.

In this example, you will retrieve a place of birth table for Bexar County, Texas, from the 2006 ACS.

- Go to the Census Bureau's Web site <www.census.gov>.
- Click on "American FactFinder."
- Under "Getting Detailed Data," select "American Community Survey – get data."
- Under "Data Sets," select "2006 American Community Survey."
- Select "Detailed Tables."
- Under "Select a geographic type," select "County."
- At the prompt, select "Texas" and then "Bexar County."
- Click "Add" to place this county in the selection box.
- Click "Next."
- Search by "Subject" (for example, search under social characteristics—birthplace, nativity, and year of entry) or "Keyword" (for example, place of birth) to find the specific table.
- Select "Table B05002. Place of Birth by Citizenship Status."
- Click "Add" to place the table in the selection box.
- Click "Show Result" to view the table (see table below).

Figure 4. **Example 1**

B05002. Place of Birth by Citizenship Status
2006 American Community Survey

	Bexar County, Texas	
	Estimate	Margin of Error
Total:	1,555,592	*****
Native:	1,370,707	+/-9,754
Born in state of residence	1,003,052	+/-13,290
Born in other state in the United States:	335,171	+/-9,407
Northeast	47,349	+/-3,308
Midwest	99,243	+/-5,649
South	101,721	+/-7,715
West	86,858	+/-6,768
Born outside the United States:	32,484	+/-3,030
Puerto Rico	7,387	+/-2,198
U.S. Island Areas	1,049	+/-551
Born abroad of American parent(s)	24,048	+/-2,372
Foreign born:	184,885	+/-9,754
Naturalized U.S. citizen	66,924	+/-4,110
Not a U.S. citizen	117,961	+/-8,600

Source: U.S. Census Bureau, 2006 American FactFinder, accessed at <http://factfinder.census.gov>.

The American FactFinder is most useful as a way to quickly look up a few ACS estimates for a few areas, but it is not practical for applications requiring large volumes of data for large numbers of small areas—such as income, education, and language data for all block groups nationwide. Users who need large quantities of ACS data may want to acquire the ACS Summary File (available for download from the Census Bureau's Web site). Like the Census 2000 Summary Files, the ACS Summary File provides a large number of data tables for a range of geographic levels.

As an alternative to the Summary File, users who know the specific tables and geographic levels they want can go to the "Download Center," which lists ACS and other data products available for download from the Census Bureau's Web site. For example, a business needing home value data for all counties could download the data to a file as described in the example below.

To download an ACS table from the Download Center, follow the steps below.

In this example, you will download a file providing home value for all counties based on the 2006 ACS.

- Go to the Census Bureau's Web site <www.census.gov>.
- Click on "American FactFinder."
- Click on "Download Center."
- Under "Select a Data Set," select "2006 American Community Survey."
- Under "Select a Geographic Summary Level," select "All Counties."
- For "Select a Download Method," use the default option—"Selected Detailed Tables."
- Click "Go."
- Search by "Subject" (for example, Housing Characteristics—Value of Home) or "Keyword" (for example, home value) to find the specific table.
- Select "Table B25075. Value for Owner-Occupied Housing Units."
- Click "Add" to place the table in the selection box.
- Click "Next."
- Click "Start Download" to initiate the download.

Downloaded files make it easier to use ACS data for large numbers of small area geographies but leave it to users to aggregate the data according to their interests. Some businesses may find it more convenient to acquire ACS data for aggregations from private suppliers, whose products are geared to dissemination for user-defined areas.

Note—ACS PUMS files are not available at the Download Center. To download ACS PUMS products, go to the ACS main page and using the Access Data tab, access the PUMS products link. Once you select the specific PUMS files of interest, you can download the PUMS data and view documentation. A page opens that prompts you to select specifications for the PUMS data to be downloaded. For more information about using PUMS data, refer to the handbook in this series for PUMS data users.

Confidentiality, Statistical Reliability, and the Reporting of ACS Data

The Census Bureau has indicated that Census 2000 long-form products can be used as a guide to the data that will be reported from the ACS and their geographic levels. Tables such as Households by Income, which Census 2000 reported for block groups, will be reported for block groups in the ACS. Tables such as Ancestry that were reported only at the census tract level will be limited to tract data in the ACS.

Confidentiality

Like the decennial census, the ACS is a mandatory survey, with response required by law. Like the census, the same law that requires response mandates that these responses be kept confidential. Individual data from the ACS are not shared with anyone outside the Census Bureau and the Census Bureau takes great care to prevent the disclosure of information on individual respondents through released ACS data products.

Similar to the census, the ACS prevents disclosure by not reporting data for individuals and households but by releasing aggregate statistics for geographic areas. The one exception is the Public Use Microdata Samples (PUMS) product, which provides a sample of individual level records but preserves confidentiality by not identifying their detailed geographic location. The individual level data reported on PUMS files are anonymous—no names or addresses are identified, and data are reported only for large Public Use Microdata Areas (PUMAs). Since PUMAs have a minimum population of 100,000, one knows little about where these respondents live, and confidentiality is well protected.

The Census Bureau also guards against the risk of disclosing individual information in aggregated data. Rare characteristics are "swapped" from one geography to another to add uncertainty. ACS block group data reflect only a sample of people living in the block group over a 5-year period, so the identity of individuals—even those with rare characteristics—is very well protected by this disclosure avoidance process.

Statistical Reliability and the Reporting of ACS Data

The ACS also limits the release of some data on the basis of expected sampling error (see discussion of sampling error below). Two versions of many of the detailed tables have been designed and are produced—a table that shows the information in full detail, and a version with fewer categories, known as a collapsed version. The 1-year and 3-year ACS tables are subject to a test of statistical reliability. For a given geographic area, if a table does not pass the reliability test, a collapsed version of the table that reports fewer categories may be provided. If the collapsed table fails to pass the reliability test, the data are not reported for that geographic area.

For example, the full Household Income table reports 16 categories of income and is illustrated in Figure 5 with estimates reflecting 2006 ACS data for the nation.

If, for a smaller area, the full table does not pass the reliability check, data would be provided only for the collapsed version of the table shown in Figure 6 (again shown with national level data) that reports only 10 categories of income—as long as this table passes the reliability check. When the full detailed table passes, the collapsed version of the table is also published.

One might expect the reliability checks to have an especially severe impact on the 5-year data, which are released for small areas, but the 5-year data are not subject to reliability tests or to collapsing or suppression. The reason is that the 5-year data are regarded as the primary ACS products for long-form replacement, and it is understood that aggregations to user-defined areas require the provision of all building-block geographies, regardless of reliability.

Collapsing and suppression may make the 1-year and 3-year data difficult for some users, as availability may not be consistent. For example, a business examining income by age of householder for a set of cities may find that the data are suppressed for some, and that only the collapsed data are available for others. Year-to-year disruptions can be expected, as an area might pass the reliability test in some years but not others. In rare cases, the data for a specific city might be provided 1 year, suppressed the next, provided in collapsed form the following year, and then released in full form the year after that.

For applications requiring the availability of all tables for all areas in all release years, users should look to the 5-year ACS data.

Figure 5. **Example of a Full Detailed Table —2006 Households by Income, U.S. Total**

	United States	
	Estimate	Margin of Error
Total:	111,617,402	+/-145,530
Less than $10,000	8,898,696	+/-49,557
$10,000 to $14,999	6,639,877	+/-39,268
$15,000 to $19,999	6,222,117	+/-38,960
$20,000 to $24,999	6,499,511	+/-39,220
$25,000 to $29,999	6,162,001	+/-38,073
$30,000 to $34,999	6,284,821	+/-40,742
$35,000 to $39,999	5,733,333	+/-38,614
$40,000 to $44,999	5,699,321	+/-38,266
$45,000 to $49,999	5,078,803	+/-33,389
$50,000 to $59,999	9,465,003	+/-46,546
$60,000 to $74,999	11,756,886	+/-50,866
$75,000 to $99,999	13,214,551	+/-53,784
$100,000 to $124,999	7,933,973	+/-46,011
$125,000 to $149,999	4,230,233	+/-29,373
$150,000 to $199,999	3,981,276	+/-28,559
$200,000 or more	3,817,000	+/-23,780

Source: U.S. Census Bureau, 2006 American FactFinder, accessed at <http://factfinder.census.gov>.

Figure 6. **Example of a Collapsed Detailed Table—2006 Households by Income, U.S. Total**

	United States	
	Estimate	Margin of Error
Total:	111,617,402	+/-145,530
Less than $10,000	8,898,696	+/-49,557
$10,000 to $14,999	6,639,877	+/-39,268
$15,000 to $24,999	12,721,628	+/-53,419
$25,000 to $34,999	12,446,822	+/-51,908
$35,000 to $49,999	16,511,457	+/-67,757
$50,000 to $74,999	21,221,889	+/-74,063
$75,000 to $99,999	13,214,551	+/-53,784
$100,000 to $149,999	12,164,206	+/-52,216
$150,000 to $199,999	3,981,276	+/-28,559
$200,000 or more	3,817,000	+/-23,780

Source: U.S. Census Bureau, 2006 American FactFinder, accessed at <http://factfinder.census.gov>.

Accuracy of ACS Data

ACS data are estimates and are subject to error and uncertainty. This is nothing new, as previous census long-form data were also estimates. However, the frequent releases with annual fluctuations, and the widespread use of ACS data make it especially important that users understand the nature and degree of error that can be involved.

Error occurs when the characteristics described by a survey differ from those of the actual population. For example, the ACS might indicate that a county has a median household income of $49,500 when the actual median is $48,200. Because the actual median is unknown, there is no way to know exactly how much the survey estimate differs from the truth, but there are ways to estimate the amount of error one could expect in ACS estimates.

All estimates, including those from the ACS, are subject to two major types of error—sampling error and nonsampling error.

> *Sampling error* results when estimates for an entire population are based on information gathered from a sample or portion of that population. Despite efforts to collect data from a sample that is representative of the whole population, sample results will always differ somewhat from those that would be obtained by interviewing everyone. Sampling error is the only source of error that can be easily quantified and thus reported in the margins of error shown in ACS data products.
>
> *Nonsampling error* results from all other sources and would exist even if information were collected from the entire population. For example, a respondent might—intentionally or unintentionally—inaccurately report their age, income, occupation, or educational attainment. Or they might provide their best guess of the year their housing unit was built. Nonsampling error also can result from nonresponse. A person who declines to report their "means of transportation to work" might have a "Drove alone" response imputed, when in fact they ride a bicycle to work. Additional error can result from the process of controlling ACS responses to reflect the Census Bureau's independent county population and housing estimates. If the Census Bureau estimates a county's population to be 127,400, but in reality it is 125,100, the error in the population estimate contributes additional error to ACS estimates. As noted earlier, measures of nonsampling error are not included in the margins of error displayed with ACS estimates. See Appendix 6 for further detail on nonsampling error.

The precise error in specific ACS estimates is never known, as it is not possible to measure the accuracy of each individual response, the error of the Census Bureau population estimates used as controls, or the error introduced by a specific sample of the population. However, there are formulas for estimating sampling error, which is larger for small samples, and would be zero if the sample included all people and households. Nonsampling error is harder to measure, and important information on nonsampling error is provided on the ACS Web site—under "Quality Measures" <http://www.census.gov/acs/www/UseData/sse/> and "Accuracy of the Data" <http://www.census.gov/acs/www/UseData/Accuracy/Accuracy1.htm>.

When considering sampling error, statisticians think not of the error of a single sample but of the distribution of errors that one would expect if we could complete many surveys based on many samples of a given size for the area in question. In general, the larger the sample, the smaller the expected sampling error. So, for a county of 100,000 households, samples of 20,000 households would produce estimates with less sampling error than samples of 10,000 households. A specific sample of 20,000 households might find that 40 percent are families with children. Because we do not know the actual percent for this county, we do not know how much the 40 percent estimate (from this sample) is in error. However, one can estimate the "standard error" expected for samples of 20,000 households out of 100,000.

Based on the "standard deviation" measure that users may recall from statistics courses, the standard error describes how often estimates based on a given sample size are likely to differ by specific amounts from the actual values they seek to measure. Again, the larger the sample, the smaller the expected error. See Appendix 3 for more detailed descriptions of sampling error and the derivation of sampling error measures.

The Census Bureau computes standard errors for all ACS estimates and reports them on ACS products as the easier to interpret "margins of error." These margins of error reflect a 90-percent confidence level and are easily translated into confidence intervals around the ACS estimates. For example, if a median estimate of $48,200 has a margin of error of plus or minus $1,000, there is a 90 percent likelihood that the actual (but unknown) median income is between $47,200 and $49,200. We would have greater confidence in this estimate than a median income estimate of $48,200 for another county where the margin of error is plus or minus $1,500.

The Census Bureau recommends that users take note of the margins of error reported with ACS data to guard against drawing erroneous conclusions. For example, before committing to build a store that sells upscale merchandise, a business might require that the metropolitan market have at least 50,000 households with incomes over $75,000 and a householder over age 55. The most recent ACS estimate might indicate 53,000 such households in a market but with a margin of error of +/- 5,000. In other words, there is a 90 percent chance that the observed estimate of 53,000 target households is within 5,000 of the actual number. The actual number could be less than 50,000, so the business must consider this level of uncertainty in deciding whether or not to establish a store in this market.

The margins of error reported with ACS data can alert users to high levels of uncertainty, identifying where they should be especially cautious in drawing conclusions. However, low margins of error do not guarantee high accuracy. In the hypothetical market described above, if the estimate of 53,000 target households was associated with a margin of error of +/- 500, a business might feel highly confident that the real number is above the 50,000 minimum they require. Users should keep in mind that unmeasured nonsampling error is another source of uncertainty.

While it is important that users keep the uncertainties of ACS data in mind, they should not despair over them. Census long-form data have always had both sampling and nonsampling error, and businesses have long put these data to effective use without knowledge of the level of error. While there are multiple sources of potential error, they are not necessarily cumulative. In some areas, these errors may offset each other to produce more accurate estimates. For example, sampling error might result in the under-estimation of a county's foreign-born population, but this error could be offset if the data are controlled to a county population estimate that is too high. But while it is important to understand that sampling and nonsampling errors can offset, it is a chance occurrence that one cannot count on.

Users might also consider if ACS period estimates based on the 2-month residence rule might impair their usefulness for applications that require estimates approximating a single point in time and reflecting census residence rules. For many applications, businesses may find these differences to be inconsequential, but for others they could be a factor for businesses to consider. Even if sampling and nonsampling error could be eliminated, users would still have to determine the extent to which ACS estimates are well-suited for their applications.

Businesses that use ACS data aggregated to custom geographic areas might seek confirmation that error levels for these aggregations are substantially lower than for the block groups from which they were built. The Census Bureau has provided formulas for calculating the margins of error for such aggregations. The formulas are reasonably straightforward, but some care is required, as there are separate formulas for data that are aggregate counts, proportions, ratios, or medians. The formulas are provided and described in Appendix 3.

Single-Year and Multiyear Period Estimates

Period Versus Point-in-Time Estimates

Although census long-form data were collected over several months of nonresponse follow-up, they have been used as approximate "snapshots" of April 1 of the census year. In contrast, the ACS provides "period estimates" reflecting data collected over periods of 1, 3, and 5 years.

Single-year estimates are based on data collected over a particular calendar year. The difference with traditional census data is illustrated by the data below, which show the monthly fluctuation in percent of the population in the labor force for a hypothetical town.

Traditional census data, focused on April 1, would report 60 percent of the town's population in the labor force. However, a 1-year period estimate averages data across all 12 months to an estimate of about 45 percent—a perfectly valid period estimate but a percentage that is not observed for any one month during the year. Actual ACS 1-year estimates are more complex

Table 4. **Period Estimates Example—Percent of Population in the Labor Force**

Month											
Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
20	20	40	60	60	60	60	60	60	50	30	20

Source: U.S. Census Bureau, Artificial Data.

than this simple average, but these hypothetical data illustrate how period estimates differ from point-in-time estimates.

ACS 3-year and 5-year estimates are produced in a similar manner. Data obtained from 3 years of interviews and 5 years of interviews are combined into a single data set and reweighted. They are not averages of the 1-year estimates that were produced for those years. These new combined estimates are controlled to the average of the Census Bureau's most recent population estimates for the years in the period.

Further descriptions of ACS single-year and multiyear estimates are provided in Appendix 1.

Transition to Period Estimates

As users transition from traditional census data to ACS period estimates, some key differences will be encountered. The ACS often presents users with a choice between 1-year, 3-year, and 5-year estimates. For small areas, only 5-year estimates are released, but for larger areas, each annual release will provide 1-year, 3-year, and 5-year estimates. For example, in 2010, there will be three sets of commuting data for San Diego County—1-year estimates for 2009, 3-year estimates reflecting 2007–2009, and 5-year estimates for the period 2005–2009. Users must decide which is the most appropriate for their application.

In making this choice, users should consider the trade-off between currency and reliability. The 1-year estimates for an area tend to have higher sampling error than the 3- and 5-year estimates for that area because they are based on only 1 year of data collection (i.e., a smaller sample). But currency is enhanced because that 1 year is the most recent year. In contrast, the 3-year and 5-year estimates for an area have larger samples and lower sampling error than the 1-year estimates for that area, but they are less current because the larger samples draw from data collected in earlier years.

Consider this trade-off as faced by a small packaged goods business whose products target households with young children. The year is 2010, and the company seeks to expand to metropolitan markets in Indiana. For the large Indianapolis market and a common characteristic such as presence of young children, 1-year data reflecting 2009 might be preferred because they are the most recent and sufficiently reliable. For smaller markets, such as Muncie, the company might want the added reliability of the 3-year estimates, even though they reflect data collected as far back as 2007. Even in Indianapolis, the company might prefer 3-year or 5-year estimates if the target is smaller populations, such as households with young children and an Asian householder.

No absolutes dictate when 1-year, 3-year, and 5-year data would be preferred, but the margins of error provided with ACS data will help our hypothetical packaged goods company weigh its options in the trade-off between currency and reliability. If the company needs to compare Indianapolis to Muncie and other potential markets, it must be careful to use consistent data. If 3-year estimates are used for Muncie and other smaller markets, they should be compared with 3-year estimates for Indianapolis. Comparisons of 1-year data for the large market with 3-year data for the smaller market, could lead to erroneous conclusions.

Users also might consider the benefits of examining all three estimates—the 1-year, 3-year, and 5-year estimates. If all three point to the same conclusion, confidence in the related business decision is increased, but if the estimates suggest different conclusions, this may signal the need for caution and further investigation. Similar to getting a second and third opinion on a medical diagnosis, cross-checking multiple ACS estimates allows thoroughness that was not an option with the one-number long-form census data.

Regardless of which sets of ACS estimates are chosen, the interpretation of period estimates can be a challenge—especially for the 3-year and 5-year multiyear estimates. The ACS provides one number for each multiyear estimate, but this one number does not reflect one point in time. A 5-year estimate of households with children (based on data from 2005 through 2009) does not reflect this characteristic for 2005, 2007, 2009, or any single year. It is probably best to think of this estimate as the average over the period from 2005 through 2009. If a characteristic is changing at a steady pace, the average for that area might be a reasonable estimate for the middle year (2007 in this example) but if the percentage was stable through 2008 and changed only in 2009, the average would not be a good estimate for 2007. One cannot generically describe 5-year estimates as reflecting the middle year of data collection and in fact should never be used to describe a particular year—only a period of time.

For example, in 2010, if a business needs to know the percent of households with children for aggregations of block groups, and the estimates have to be for the year 2009, the ACS will not provide such estimates. Aggregations of block groups require 5-year estimates, which at that time, will reflect data collected from 2005–2009. However, if the business needs to know which of these market areas have had—in recent years—the highest percent of households with children, the ACS will answer that question very well.

See Appendix 1 for further guidance on the use of ACS single-year and multiyear data.

Tracking Change With ACS estimates

Tracking change with ACS data requires care. Changes in 1-year estimates are relatively straightforward, but users need to watch margins of error closely to distinguish actual change from sampling variability or "noise." For example, if the 2006 ACS reports that a county had a median household income of $53,000 (with a margin of error of +/- $1,200) and the 2007 ACS reports a median of $51,900 in 2007 (margin of error +/- $1,100), is the county really experiencing a decline in income, or might the decline be a function of sampling variability? The ACS estimate alone does not provide a definitive answer, but the margins of error can be used to make this determination. To assist users in tracking change, the Census Bureau will provide Comparison Profiles for all areas receiving 1-year estimates. Beginning with the 2007 ACS products, the Comparison Profile will provide data for multiple years (one column per year), with asterisks identifying changes that are statistically significant at the 90-percent confidence level.

Tracking change with multiyear estimates is more complicated because the data overlap for successive years. For example, 5-year estimates released in 2010 will reflect data collected from 2005 though 2009, and 5-year estimates released the following year will reflect data collected from 2006 through 2010. Both estimates share data collected in 2006, 2007, 2008, and 2009. Put another way, only about 1/5, or 20 percent, of the data in the "new" estimate is really new.

For this reason, change is best-measured using estimates reflecting nonoverlapping time periods. For example, 5-year estimates based on data collected during the period 2005 through 2009 (and released in 2010) are best compared with 5-year estimates based on data collected during the period 2010 through 2014 (and released in 2015).

Many businesses will not want to wait 5 years to measure change based on 5-year estimates. Such businesses should keep in mind that census long-form data required a wait of 10 years, and that once the initial pair of nonoverlapping 5-year ACS estimates becomes available (when the 2015 release can be compared with the 2010), a new pair will follow every year. Five-year estimates released in 2016 can be compared with 2011; those released in 2017 can be compared with 2012, and so on.

One could argue that year-to-year differences in multiyear estimates provide potentially useful information, and the Census Bureau acknowledges that a series of overlapping estimates can smooth out short-term fluctuations that can obscure longer-term trends. The key is interpreting the data correctly and with realistic expectations. For example, a business might watch annual changes in 5-year data on the Spanish-speaking population to see which of its trade areas are stable, and which are growing with respect to this characteristic. Understanding that these are annual changes in 5-year averages, the business realizes that it is not pinpointing exact rates of change but rather getting valuable information on the overall pattern of change that the census long-form data could not provide.

The complications described above illustrate why users will have to take care in transitioning to ACS data. But while ACS data require that users become knowledgeable about some issues, the payoff is frequent updates to a wide range of information. In the absence of the ACS, the hypothetical companies described above would have had to rely on 10-year old data from Census 2000, and patterns of change limited to the 10-year span between censuses.

Appendix 4 provides additional detail on comparisons of ACS data—both with other data sources and across time periods.

ACS Data and Business Applications

The ACS data are intended to take the place of census long-form data, so the suitability of ACS data for business applications is similar to the suitability of census long-form data.

Businesses with occasional or limited data needs will be well served by the prepackaged ACS data products on the Census Bureau Web site, while large scale users may prefer to acquire ACS summary files or download detailed tables. Users also can expect to acquire ACS data in value-added form from private suppliers.

For many business applications, the focus is not on the ACS data themselves but on marketing products built from them. For example, a business may be less interested in the ACS estimates of income and age of householder than in the estimates of sales potential for products (such as consumer electronics) that are derived from them. Whatever the channel for acquiring ACS data, and whether the focus is on ACS data or derived products, business users should be aware of the nature of ACS data.

The fact that the ACS has a smaller sample and larger sampling error than the Census 2000 long-form sample is important to business users, who typically need data for very small areas. The suitability of ACS data for these businesses depends on the application and the level of accuracy required. An important distinction is between applications that use small area data as building blocks to larger areas and those that focus on the characteristics of a specific small area.

Using ACS Data as Building Blocks for Larger Geographic Areas

Applications that use small areas as building blocks to business-specific trade areas should be well served by ACS data. Companies may use data supplied by private organizations to examine demographic data for custom areas such as 3-mile radii around store locations, 20-minute drive times to bank branches, or telecommunications service areas. The private suppliers' systems build aggregations from block groups, and even when ACS estimates for individual block groups have large errors, aggregations to larger areas will reduce sampling error to lower levels.

Using ACS Data for Small Geographic Areas

While aggregations have always been the best way to use small area census data (including long-form data), some businesses have used data for individual block groups or defined markets consisting of one or a few block groups. For example, a fast food business might consider the size and characteristics of the population in a 3-block group area to be critical to the success of a location. However, ACS age, income, and employment data may be too uncertain to support a decision to establish a franchise at that location. Indeed, census long-form data may have had too much error for this application as well. But with the ACS, the user can examine the margins of error reported for the three block groups (and calculate the margin of error for the aggregation) to gauge the level of uncertainty involved.

There also is a history of businesses using block group characteristics as a proxy for data on individual consumers. A publisher might have the name, address, and billing history of current subscribers to one magazine but need to append income, presence of children, and education data to these subscriber records to target prospects for a new publication. Because each subscriber's address corresponds to a block group, the publisher might infer these characteristics on a probability basis based on the characteristics of the block groups where subscribers live. This approach is error prone even if block group data are perfectly accurate, and the error associated with census or ACS data adds considerable uncertainty. For these reasons, many businesses tap the resources of household level databases that are available from commercial suppliers. These databases have their own errors and often limited content, so the temptation to use the characteristics of individual block groups persists.

From a statistical perspective, applications of this type appear reckless. The ACS might indicate that 58 percent of households in a block group have a child when the (unknown) reality is only 29 percent. Clearly, it would be a mistake to market to this block group as one with many children. But marketing efforts are not confined to a single block group, and a business might conclude that, across thousands of block groups, inferences of this type might work well enough often enough to outperform alternatives such as targeting based on county data or untargeted blanket mailings. It is up to individual businesses to know the requirements of their applications and the limitations of ACS data before committing to applications of this type.

Using Data on ZIP Codes

ZIP Code data are widely used by businesses and merit a special note. First, even though the Census Bureau plans to provide ACS data for ZIP Code Tabulation Areas (see the description of ZCTAs above), annual updates to ZCTA definitions are said to be unlikely. For this reason, ZIP Code estimates produced by private suppliers are likely to have a strong following. Second, while some proprietary ZIP Code data are aggregations from household level records (such as customer databases), private suppliers will have to aggregate ACS data to ZIP Codes from the block groups for which they are reported. Most ZIP Codes will benefit from the reduced error of block group aggregation, but some are very small, and will have high levels of sampling error.

Using Data on Population and Housing Counts

Many businesses, such as commercial developers, seek areas of rapid population growth for their untapped potential and attractive demographics. Such users need to understand that the ACS is not about population counts but rather the characteristics of the population, households, and housing.

At the county level and above, ACS population and housing totals are forced to agree with the Census Bureau's population and housing estimates. Below the county level, the Census Bureau's Master Address File (MAF) determines the distribution of sampled addresses but not the final housing or population totals. For example, the MAF count might reflect a block group's recent growth from 200 addresses in 2005 to 2,000 addresses in 2009, but the 5-year estimate will reflect addresses sampled over the 5-year period, including years when the address count was much lower. The resulting housing and population totals are further decreased or increased as they are weighted for consistency with the Census Bureau's county population and housing estimates.

ACS housing and population totals derive less from the survey itself, than from the Census Bureau's Population Estimates Program since the Census Bureau's estimates of population by race/ethnicity and age are used as controls. For data on housing and population totals, or for age, sex, race, and Hispanic ethnicity, the Census Bureau recommends using data from the Census Bureau's Population Estimates Program.

Implications for Private Sector Value Added Products

Since the 1970s, a number of firms have functioned as census data intermediaries—processing and providing data in value-added forms convenient to business users. These data companies often provide more than census data, but for decades, their census-based products have operated on a 10-year cycle coinciding with the decennial census. With each new census, they reconfigure their data reporting, consumer segmentation, mapping, and other products to incorporate the new census data and geography. Private data sources and industry-specific geography are always changing, but census data and geography have provided a fixed foundation for a 10-year period.

The transition to the ACS will affect the data industry's cycle for census-based products. The ACS now will update demographic, socioeconomic, and housing characteristics every year. It remains to be seen exactly how private suppliers will incorporate the torrent of new data into their product offerings.

It also remains to be seen if, in reporting ACS data, the data companies will include the margins of error. If there is demand among business users, there will be incentive to provide margins of error, but for some providers, the massive increase in database size (providing another number with every number) may be an obstacle. Private companies must decide whether to report the ACS's published margins of error and whether to build into their products the ability to calculate margins of error for ACS data that users may aggregate to custom geographic areas.

The private data companies produce small area estimates and projections of population and household totals, as well as characteristics, such as age, race, and income. They produce these estimates for all block groups nationwide, once per year, and in some cases, once per quarter. But as ambitious as the suppliers' efforts have been, they have not typically attempted estimates of characteristics such as education, occupation, language, disability, commuting, and many others formerly provided by the census long form. With the ACS releasing new block-group level data for these items every year, private suppliers will be challenged to incorporate them into their products. In doing so, they will have to decide whether to provide the 5-year period estimates "as is," or attempt to project the data forward to something reflecting more conventional "current year" point-in-time estimates.

The suppliers also will have to consider the impact of the ACS on the update schedule for neighborhood level segmentation systems. Although somewhat overshadowed in recent years by segmentation systems based on household-level proprietary sources, the "geodemographic" systems remain popular for some applications. These systems assign block groups to "lifestyle clusters" defined on block group census data. The assignment of block groups to the lifestyle segments can change during the decade based on private data sources, or estimates, but the segments themselves are typically defined once a decade following the release of new census data. With the annual release of ACS characteristics data for block groups, data companies can reconsider the frequency with which it makes sense to update geodemographic segmentation products.

Glossary

Accuracy. One of four key dimensions of survey quality. Accuracy refers to the difference between the survey estimate and the true (unknown) value. Attributes are measured in terms of sources of error (for example, coverage, sampling, nonresponse, measurement, and processing).

American Community Survey Alert. This periodic electronic newsletter informs data users and other interested parties about news, events, data releases, congressional actions, and other developments associated with the ACS. See <http://www.census.gov/acs/www/Special/Alerts/Latest.htm>.

American FactFinder (AFF). An electronic system for access to and dissemination of Census Bureau data on the Internet. AFF offers prepackaged data products and user-selected data tables and maps from Census 2000, the 1990 Census of Population and Housing, the 1997 and 2002 Economic Censuses, the Population Estimates Program, annual economic surveys, and the ACS.

Block group. A subdivision of a census tract (or, prior to 2000, a block numbering area), a block group is a cluster of blocks having the same first digit of their four-digit identifying number within a census tract.

Census geography. A collective term referring to the types of geographic areas used by the Census Bureau in its data collection and tabulation operations, including their structure, designations, and relationships to one another. See <http://www.census.gov/geo/www/index.html>.

Census tract. A small, relatively permanent statistical subdivision of a county delineated by a local committee of census data users for the purpose of presenting data. Census tract boundaries normally follow visible features, but may follow governmental unit boundaries and other nonvisible features; they always nest within counties. Designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions at the time of establishment, census tracts average about 4,000 inhabitants.

Coefficient of variation (CV). The ratio of the standard error (square root of the variance) to the value being estimated, usually expressed in terms of a percentage (also known as the relative standard deviation). The lower the CV, the higher the relative reliability of the estimate.

Comparison profile. Comparison profiles are available from the American Community Survey for 1-year estimates beginning in 2007. These tables are available for the U.S., the 50 states, the District of Columbia, and geographic areas with a population of more than 65,000.

Confidence interval. The sample estimate and its standard error permit the construction of a confidence interval that represents the degree of uncertainty about the estimate. A 90-percent confidence interval can be interpreted roughly as providing 90 percent certainty that the interval defined by the upper and lower bounds contains the true value of the characteristic.

Confidentiality. The guarantee made by law (Title 13, United States Code) to individuals who provide census information, regarding nondisclosure of that information to others.

Consumer Price Index (CPI). The CPI program of the Bureau of Labor Statistics produces monthly data on changes in the prices paid by urban consumers for a representative basket of goods and services.

Controlled. During the ACS weighting process, the intercensal population and housing estimates are used as survey controls. Weights are adjusted so that ACS estimates conform to these controls.

Current Population Survey (CPS). The CPS is a monthly survey of about 50,000 households conducted by the Census Bureau for the Bureau of Labor Statistics. The CPS is the primary source of information on the labor force characteristics of the U.S. population.

Current residence. The concept used in the ACS to determine who should be considered a resident of a sample address. Everyone who is currently living or staying at a sample address is considered a resident of that address, except people staying there for 2 months or less. People who have established residence at the sample unit and are away for only a short period of time are also considered to be current residents.

Custom tabulations. The Census Bureau offers a wide variety of general purpose data products from the American Community Survey (ACS). These products are designed to meet the needs of the majority of data users and contain predefined sets of data for standard census geographic areas, including both political and statistical geography. These products are available on the American FactFinder and the ACS Web site.

For users with data needs not met through the general purpose products, the Census Bureau offers "custom" tabulations on a cost-reimbursable basis, with the American Community Survey Custom Tabulation program. Custom tabulations are created by tabulating data from ACS microdata files. They vary in size, complexity, and cost depending on the needs of the sponsoring client.

Data profiles. Detailed tables that provide summaries by social, economic, and housing characteristics. There is a new ACS demographic and housing units profile that should be used if official estimates from the Population Estimates Program are not available.

Detailed tables. Approximately 1,200 different tables that contain basic distributions of characteristics. These tables provide the most detailed data and are the basis for other ACS products.

Disclosure avoidance (DA). Statistical methods used in the tabulation of data prior to releasing data products to ensure the confidentiality of responses. See Confidentiality.

Estimates. Numerical values obtained from a statistical sample and assigned to a population parameter. Data produced from the ACS interviews are collected from samples of housing units. These data are used to produce estimates of the actual figures that would have been obtained by interviewing the entire population using the same methodology.

Five-year estimates. Estimates based on 5 years of ACS data. These estimates reflect the characteristics of a geographic area over the entire 5-year period and will be published for all geographic areas down to the census block group level.

File Transfer Protocol (FTP) site. A Web site that allows data files to be downloaded from the Census Bureau Web site.

Geographic comparison tables. More than 80 single-variable tables comparing key indicators for geographies other than states.

Geographic summary level. A geographic summary level specifies the content and the hierarchical relationships of the geographic elements that are required to tabulate and summarize data. For example, the county summary level specifies the state-county hierarchy. Thus, both the state code and the county code are required to uniquely identify a county in the United States or Puerto Rico.

Group quarters (GQ) facilities. A GQ facility is a place where people live or stay that is normally owned or managed by an entity or organization providing housing and/or services for the residents. These services may include custodial or medical care, as well as other types of assistance. Residency is commonly restricted to those receiving these services. People living in GQ facilities are usually not related to each other. The ACS collects data from people living in both housing units and GQ facilities.

Group quarters (GQ) population. The number of persons residing in GQ facilities.

Item allocation rates. Allocation is a method of imputation used when values for missing or inconsistent items cannot be derived from the existing response record. In these cases, the imputation must be based on other techniques such as using answers from other people in the household, other responding housing units, or people believed to have similar characteristics. Such donors are reflected in a table referred to as an allocation matrix. The rate is percentage of times this method is used.

Margin of error (MOE). Some ACS products provide an MOE instead of confidence intervals. An MOE is the difference between an estimate and its upper or lower confidence bounds. Confidence bounds can be created by adding the margin of error to the estimate (for the upper bound) and subtracting the margin of error from the estimate (for the lower bound). All published ACS margins of error are based on a 90-percent confidence level.

Multiyear estimates. Three- and five-year estimates based on multiple years of ACS data. Three-year estimates will be published for geographic areas with a population of 20,000 or more. Five-year estimates will be published for all geographic areas down to the census block group level.

Narrative profile. A data product that includes easy-to-read descriptions for a particular geography.

Nonsampling error. Total survey error can be classified into two categories—sampling error and nonsampling error. Nonsampling error includes measurement errors due to interviewers, respondents, instruments, and mode; nonresponse error; coverage error; and processing error.

Period estimates. An estimate based on information collected over a period of time. For ACS the period is either 1 year, 3 years, or 5 years.

Point-in-time estimates. An estimate based on one point in time. The decennial census long-form estimates for Census 2000 were based on information collected as of April 1, 2000.

Population estimates program. Official Census Bureau estimates of the population of the United States, states, metropolitan areas, cities and towns, and counties; also official Census Bureau estimates of housing units (HUs).

Public Use Microdata Area (PUMA). An area that defines the extent of territory for which the Census Bureau releases Public Use Microdata Sample (PUMS) records.

Public Use Microdata Sample (PUMS) files. Computerized files that contain a sample of individual records, with identifying information removed, showing the population and housing characteristics of the units, and people included on those forms.

Puerto Rico Community Survey (PRCS). The counterpart to the ACS that is conducted in Puerto Rico.

Quality measures. Statistics that provide information about the quality of the ACS data. The ACS releases four different quality measures with the annual data release: 1) initial sample size and final interviews; 2) coverage rates; 3) response rates, and; 4) item allocation rates for all collected variables. The ACS Quality Measures Web site provides these statistics each year. In addition, the coverage rates are also available for males and females separately.

Reference period. Time interval to which survey responses refer. For example, many ACS questions refer to the day of the interview; others refer to "the past 12 months" or "last week."

Residence rules. The series of rules that define who (if anyone) is considered to be a resident of a sample address for purposes of the survey or census.

Sampling error. Errors that occur because only part of the population is directly contacted. With any sample, differences are likely to exist between the characteristics of the sampled population and the larger group from which the sample was chosen.

Sampling variability. Variation that occurs by chance because a sample is surveyed rather than the entire population.

Selected population profiles. An ACS data product that provides certain characteristics for a specific race or ethnic group (for example, Alaska Natives) or other population subgroup (for example, people aged 60 years and over). This data product is produced directly from the sample microdata (that is, not a derived product).

Single-year estimates. Estimates based on the set of ACS interviews conducted from January through December of a given calendar year. These estimates are published each year for geographic areas with a population of 65,000 or more.

Standard error. The standard error is a measure of the deviation of a sample estimate from the average of all possible samples.

Statistical significance. The determination of whether the difference between two estimates is not likely to be from random chance (sampling error) alone. This determination is based on both the estimates themselves and their standard errors. For ACS data, two estimates are "significantly different at the 90 percent level" if their difference is large enough to infer that there was a less than 10 percent chance that the difference came entirely from random variation.

Subject tables. Data products organized by subject area that present an overview of the information that analysts most often receive requests for from data users.

Summary files. Consist of detailed tables of Census 2000 social, economic, and housing characteristics compiled from a sample of approximately 19 million housing units (about 1 in 6 households) that received the Census 2000 long-form questionnaire.

Thematic maps. Display geographic variation in map format from the geographic ranking tables.

Three-year estimates. Estimates based on 3 years of ACS data. These estimates are meant to reflect the characteristics of a geographic area over the entire 3-year period. These estimates will be published for geographic areas with a population of 20,000 or more.

Appendix 1.

Understanding and Using ACS Single-Year and Multiyear Estimates

What Are Single-Year and Multiyear Estimates?

Understanding Period Estimates

The ACS produces period estimates of socioeconomic and housing characteristics. It is designed to provide estimates that describe the average characteristics of an area over a specific time period. In the case of ACS single-year estimates, the period is the calendar year (e.g., the 2007 ACS covers January through December 2007). In the case of ACS multiyear estimates, the period is either 3 or 5 calendar years (e.g., the 2005–2007 ACS estimates cover January 2005 through December 2007, and the 2006–2010 ACS estimates cover January 2006 through December 2010). The ACS multiyear estimates are similar in many ways to the ACS single-year estimates, however they encompass a longer time period. As discussed later in this appendix, the differences in time periods between single-year and multiyear ACS estimates affect decisions about which set of estimates should be used for a particular analysis.

While one may think of these estimates as representing average characteristics over a single calendar year or multiple calendar years, it must be remembered that the 1-year estimates are not calculated as an average of 12 monthly values and the multiyear estimates are not calculated as the average of either 36 or 60 monthly values. Nor are the multiyear estimates calculated as the average of 3 or 5 single-year estimates. Rather, the ACS collects survey information continuously nearly every day of the year and then aggregates the results over a specific time period—1 year, 3 years, or 5 years. The data collection is spread evenly across the entire period represented so as not to over-represent any particular month or year within the period.

Because ACS estimates provide information about the characteristics of the population and housing for areas over an entire time frame, ACS single-year and multiyear estimates contrast with "point-in-time" estimates, such as those from the decennial census long-form samples or monthly employment estimates from the Current Population Survey (CPS), which are designed to measure characteristics as of a certain date or narrow time period. For example, Census 2000 was designed to measure the characteristics of the population and housing in the United States based upon data collected around April 1, 2000, and thus its data reflect a narrower time frame than ACS data. The monthly CPS collects data for an even narrower time frame, the week containing the 12th of each month.

Implications of Period Estimates

Most areas have consistent population characteristics throughout the calendar year, and their period estimates may not look much different from estimates that would be obtained from a "point-in-time" survey design. However, some areas may experience changes in the estimated characteristics of the population, depending on when in the calendar year measurement occurred. For these areas, the ACS period estimates (even for a single-year) may noticeably differ from "point-in-time" estimates. The impact will be more noticeable in smaller areas where changes such as a factory closing can have a large impact on population characteristics, and in areas with a large physical event such as Hurricane Katrina's impact on the New Orleans area. This logic can be extended to better interpret 3-year and 5-year estimates where the periods involved are much longer. If, over the full period of time (for example, 36 months) there have been major or consistent changes in certain population or housing characteristics for an area, a period estimate for that area could differ markedly from estimates based on a "point-in-time" survey.

An extreme illustration of how the single-year estimate could differ from a "point-in-time" estimate within the year is provided in Table 1. Imagine a town on the Gulf of Mexico whose population is dominated by retirees in the winter months and by locals in the summer months. While the percentage of the population in the labor force across the entire year is about 45 percent (similar in concept to a period estimate), a "point-in-time" estimate for any particular month would yield estimates ranging from 20 percent to 60 percent.

Table 1. **Percent in Labor Force—Winter Village**

Month											
Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
20	20	40	60	60	60	60	60	60	50	30	20

Source: U.S. Census Bureau, Artificial Data.

The important thing to keep in mind is that ACS single-year estimates describe the population and characteristics of an area for the full year, not for any specific day or period within the year, while ACS multiyear estimates describe the population and characteristics of an area for the full 3- or 5-year period, not for any specific day, period, or year within the multiyear time period.

Release of Single-Year and Multiyear Estimates

The Census Bureau has released single-year estimates from the full ACS sample beginning with data from the 2005 ACS. ACS 1-year estimates are published annually for geographic areas with populations of 65,000 or more. Beginning in 2008 and encompassing 2005–2007, the Census Bureau will publish annual ACS 3-year estimates for geographic areas with populations of 20,000 or more. Beginning in 2010, the Census Bureau will release ACS 5-year estimates (encompassing 2005–2009) for all geographic areas —down to the tract and block group levels. While eventually all three data series will be available each year, the ACS must collect 5 years of sample before that final set of estimates can be released. This means that in 2008 only 1-year and 3-year estimates are available for use, which means that data are only available for areas with populations of 20,000 and greater.

New issues will arise when multiple sets of multiyear estimates are released. The multiyear estimates released in consecutive years consist mostly of overlapping years and shared data. As shown in Table 2, consecutive 3-year estimates contain 2 years of overlapping coverage (for example, the 2005–2007 ACS estimates share 2006 and 2007 sample data with the 2006–2008 ACS estimates) and consecutive 5-year estimates contain 4 years of overlapping coverage.

Table 2. **Sets of Sample Cases Used in Producing ACS Multiyear Estimates**

Type of estimate	Year of Data Release				
	2008	2009	2010	2011	2012
	Years of Data Collection				
3-year estimates	2005–2007	2006–2008	2007–2009	2008–2010	2009–2011
5-year estimates	Not Available	Not Available	2005–2009	2006–2010	2007–2011

Source: U.S. Census Bureau.

Differences Between Single-Year and Multiyear ACS Estimates

Currency

Single-year estimates provide more current information about areas that have changing population and/or housing characteristics because they are based on the most current data—data from the past year. In contrast, multiyear estimates provide less current information because they are based on both data from the previous year and data that are 2 and 3 years old. As noted earlier, for many areas with minimal change taking place, using the "less current" sample used to produce the multiyear estimates may not have a substantial influence on the estimates. However, in areas experiencing major changes over a given time period, the multiyear estimates may be quite different from the single-year estimates for any of the individual years. Single-year and multiyear estimates are not expected to be the same because they are based on data from two different time periods. This will be true even if the ACS single year is the midyear of the ACS multiyear period (e.g., 2007 single year, 2006–2008 multiyear).

For example, suppose an area has a growing Hispanic population and is interested in measuring the percent of the population who speak Spanish at home. Table 3 shows a hypothetical set of 1-year and 3-year estimates. Comparing data by release year shows that for an area such as this with steady growth, the 3-year estimates for a period are seen to lag behind the estimates for the individual years.

Reliability

Multiyear estimates are based on larger sample sizes and will therefore be more reliable. The 3-year estimates are based on three times as many sample cases as the 1-year estimates. For some characteristics this increased sample is needed for the estimates to be reliable enough for use in certain applications. For other characteristics the increased sample may not be necessary.

Table 3. **Example of Differences in Single- and Multiyear Estimates—Percent of Population Who Speak Spanish at Home**

Year of data release	1-year estimates		3-year estimates	
	Time period	Estimate	Time period	Estimate
2003	2002	13.7	2000–2002	13.4
2004	2003	15.1	2001–2003	14.4
2005	2004	15.9	2002–2004	14.9
2006	2005	16.8	2003–2005	15.9

Source: U.S. Census Bureau, Artificial Data.

Multiyear estimates are the only type of estimates available for geographic areas with populations of less than 65,000. Users may think that they only need to use multiyear estimates when they are working with small areas, but this isn't the case. Estimates for large geographic areas benefit from the increased sample resulting in more precise estimates of population and housing characteristics, especially for subpopulations within those areas.

In addition, users may determine that they want to use single-year estimates, despite their reduced reliability, as building blocks to produce estimates for meaningful higher levels of geography. These aggregations will similarly benefit from the increased sample sizes and gain reliability.

Deciding Which ACS Estimate to Use

Three primary uses of ACS estimates are to understand the characteristics of the population of an area for local planning needs, make comparisons across areas, and assess change over time in an area. Local planning could include making local decisions such as where to locate schools or hospitals, determining the need for services or new businesses, and carrying out transportation or other infrastructure analysis. In the past, decennial census sample data provided the most comprehensive information. However, the currency of those data suffered through the intercensal period, and the ability to assess change over time was limited. ACS estimates greatly improve the currency of data for understanding the characteristics of housing and population and enhance the ability to assess change over time.

Several key factors can guide users trying to decide whether to use single-year or multiyear ACS estimates for areas where both are available: intended use of the estimates, precision of the estimates, and currency of the estimates. All of these factors, along with an understanding of the differences between single-year and multiyear ACS estimates, should be taken into consideration when deciding which set of estimates to use.

Understanding Characteristics

For users interested in obtaining estimates for small geographic areas, multiyear ACS estimates will be the only option. For the very smallest of these areas (less than 20,000 population), the only option will be to use the 5-year ACS estimates. Users have a choice of two sets of multiyear estimates when analyzing data for small geographic areas with populations of at least 20,000. Both 3-year and 5-year ACS estimates will be available. Only the largest areas with populations of 65,000 and more receive all three data series.

The key trade-off to be made in deciding whether to use single-year or multiyear estimates is between currency and precision. In general, the single-year estimates are preferred, as they will be more relevant to the current conditions. However, the user must take into account the level of uncertainty present in the single-year estimates, which may be large for small subpopulation groups and rare characteristics. While single-year estimates offer more current estimates, they also have higher sampling variability. One measure, the coefficient of variation (CV) can help you determine the fitness for use of a single-year estimate in order to assess if you should opt instead to use the multiyear estimate (or if you should use a 5-year estimate rather than a 3-year estimate). The CV is calculated as the ratio of the standard error of the estimate to the estimate, times 100. A single-year estimate with a small CV is usually preferable to a multiyear estimate as it is more up to date. However, multiyear estimates are an alternative option when a single-year estimate has an unacceptably high CV.

Table 4 illustrates how to assess the reliability of 1-year estimates in order to determine if they should be used. The table shows the percentage of households where Spanish is spoken at home for ACS test counties Broward, Florida, and Lake, Illinois. The standard errors and CVs associated with those estimates are also shown.

In this illustration, the CV for the single-year estimate in Broward County is 1.0 percent (0.2/19.9) and in Lake County is 1.3 percent (0.2/15.9). Both are sufficiently small to allow use of the more current single-year estimates.

Single-year estimates for small subpopulations (e.g., families with a female householder, no husband, and related children less than 18 years) will typically have larger CVs. In general, multiyear estimates are preferable to single-year estimates when looking at estimates for small subpopulations.

For example, consider Sevier County, Tennessee, which had an estimated population of 76,632 in 2004 according to the Population Estimates Program. This population is larger than the Census Bureau's 65,000-population requirement for publishing 1-year estimates. However, many subpopulations within this geographic area will be much smaller than 65,000. Table 5 shows an estimated 21,881 families in Sevier County based on the 2000–2004 multiyear estimate; but only 1,883 families with a female householder, no husband present, with related children under 18 years. Not surprisingly, the 2004 ACS estimate of the poverty rate (38.3 percent) for this subpopulation has a large standard error (SE) of 13.0 percentage points. Using this information we can determine that the CV is 33.9 percent (13.0/38.3).

For such small subpopulations, users obtain more precision using the 3-year or 5-year estimate. In this example, the 5-year estimate of 40.2 percent has an SE of 4.9 percentage points that yields a CV of 12.2 percent (4.9/40.2), and the 3-year estimate of 40.4 percent has an SE of 6.8 percentage points which yields a CV of 16.8 percent (6.8/40.4).

Users should think of the CV associated with an estimate as a way to assess "fitness for use." The CV threshold that an individual should use will vary based on the application. In practice there will be many estimates with CVs over desirable levels. A general guideline when working with ACS estimates is that, while data are available at low geographic levels, in situations where the CVs for these estimates are high, the reliability of the estimates will be improved by aggregating such estimates to a higher geographic level. Similarly, collapsing characteristic detail (for example, combining individual age categories into broader categories) can allow you to improve the reliability of the aggregate estimate, bringing the CVs to a more acceptable level.

Table 4. **Example of How to Assess the Reliability of Estimates—Percent of Population Who Speak Spanish at Home**

County	Estimate	Standard error	Coefficient of variation
Broward County, FL	19.9	0.2	1.0
Lake County, IL	15.9	0.2	1.3

Source: U.S. Census Bureau, Multiyear Estimates Study data.

Table 5. **Percent in Poverty by Family Type for Sevier County, TN**

	2000–2004	2000–2004		2002–2004		2004	
	Total family type	Pct. in poverty	SE	Pct. in poverty	SE	Pct. in poverty	SE
All families	21,881	9.5	0.8	9.7	1.3	10.0	2.3
With related children under 18 years	9,067	15.3	1.5	16.5	2.4	17.8	4.5
Married-couple families	17,320	5.8	0.7	5.4	0.9	7.9	2.0
With related children under 18 years	6,633	7.7	1.2	7.3	1.7	12.1	3.9
Families with female householder, no husband	3,433	27.2	3.0	26.7	4.8	19.0	7.2
With related children under 18 years	1,883	40.2	4.9	40.4	6.8	38.3	13.0

Source: U.S. Census Bureau, Multiyear Estimates Study data.

Making Comparisons

Often users want to compare the characteristics of one area to those of another area. These comparisons can be in the form of rankings or of specific pairs of comparisons. Whenever you want to make a comparison between two different geographic areas you need to take the type of estimate into account. It is important that comparisons be made within the same estimate type. That is, 1-year estimates should only be compared with other 1-year estimates, 3-year estimates should only be compared with other 3-year estimates, and 5-year estimates should only be compared with other 5-year estimates.

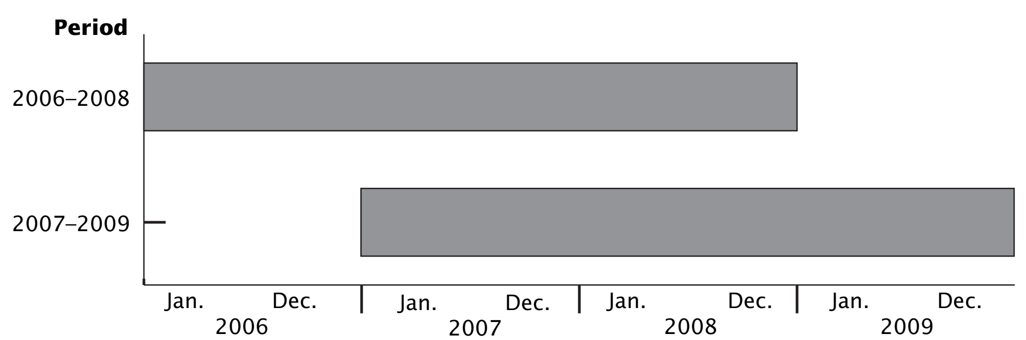
You certainly can compare characteristics for areas with populations of 30,000 to areas with populations of 100,000 but you should use the data set that they have in common. In this example you could use the 3-year or the 5-year estimates because they are available for areas of 30,000 and areas of 100,000.

Assessing Change

Users are encouraged to make comparisons between sequential single-year estimates. Specific guidance on making these comparisons and interpreting the results are provided in Appendix 4. Starting with the 2007 ACS, a new data product called the comparison profile will do much of the statistical work to identify statistically significant differences between the 2007 ACS and the 2006 ACS.

As noted earlier, caution is needed when using multiyear estimates for estimating year-to-year change in a particular characteristic. This is because roughly two-thirds of the data in a 3-year estimate overlap with the data in the next year's 3-year estimate (the overlap is roughly four-fifths for 5-year estimates). Thus, as shown in Figure 1, when comparing 2006–2008 3-year estimates with 2007–2009 3-year estimates, the differences in overlapping multiyear estimates are driven by differences in the nonoverlapping years. A data user interested in comparing 2009 with 2008 will not be able to isolate those differences using these two successive 3-year estimates. Figure 1 shows that the difference in these two estimates describes the difference between 2009 and 2006. While the interpretation of this difference is difficult, these comparisons can be made with caution. Users who are interested in comparing overlapping multiyear period estimates should refer to Appendix 4 for more information.

Figure 1. **Data Collection Periods for 3–Year Estimates**

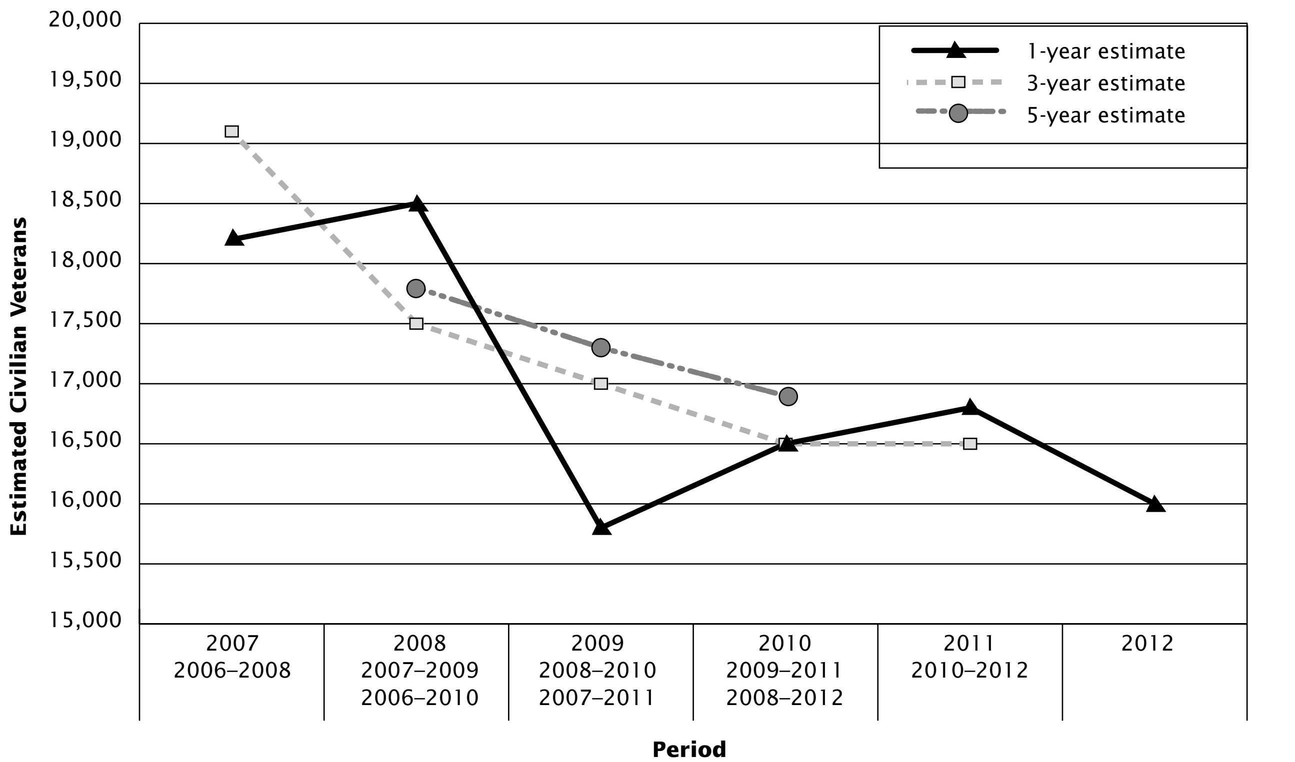


Source: U.S. Census Bureau.

Variability in single-year estimates for smaller areas (near the 65,000-publication threshold) and small subgroups within even large areas may limit the ability to examine trends. For example, single-year estimates for a characteristic with a high CV vary from year to year because of sampling variation obscuring an underlying trend. In this case, multiyear estimates may be useful for assessing an underlying, long-term trend. Here again, however, it must be recognized that because the multiyear estimates have an inherent smoothing, they will tend to mask rapidly developing changes. Plotting the multiyear estimates as representing the middle year is a useful tool to illustrate the smoothing effect of the multiyear weighting methodology. It also can be used to assess the "lagging effect" in the multiyear estimates. As a general rule, users should not consider a multiyear estimate as a proxy for the middle year of the period. However, this could be the case under some specific conditions, as is the case when an area is experiencing growth in a linear trend.

As Figure 2 shows, while the single-year estimates fluctuate from year to year without showing a smooth trend, the multiyear estimates, which incorporate data from multiple years, evidence a much smoother trend across time.

Figure 2. **Civilian Veterans, County X Single-Year, Multiyear Estimates**



Source: U.S. Census Bureau. Based on data from the Multiyear Estimates Study.

Summary of Guidelines

Multiyear estimates should, in general, be used when single-year estimates have large CVs or when the precision of the estimates is more important than the currency of the data. Multiyear estimates should also be used when analyzing data for smaller geographies and smaller populations in larger geographies. Multiyear estimates are also of value when examining change over nonoverlapping time periods and for smoothing data trends over time.

Single-year estimates should, in general, be used for larger geographies and populations when currency is more important than the precision of the estimates. Single-year estimates should be used to examine year-to-year change for estimates with small CVs. Given the availability of a single-year estimate, calculating the CV provides useful information to determine if the single-year estimate should be used. For areas believed to be experiencing rapid changes in a characteristic, single-year estimates should generally be used rather than multiyear estimates as long as the CV for the single-year estimate is reasonable for the specific usage.

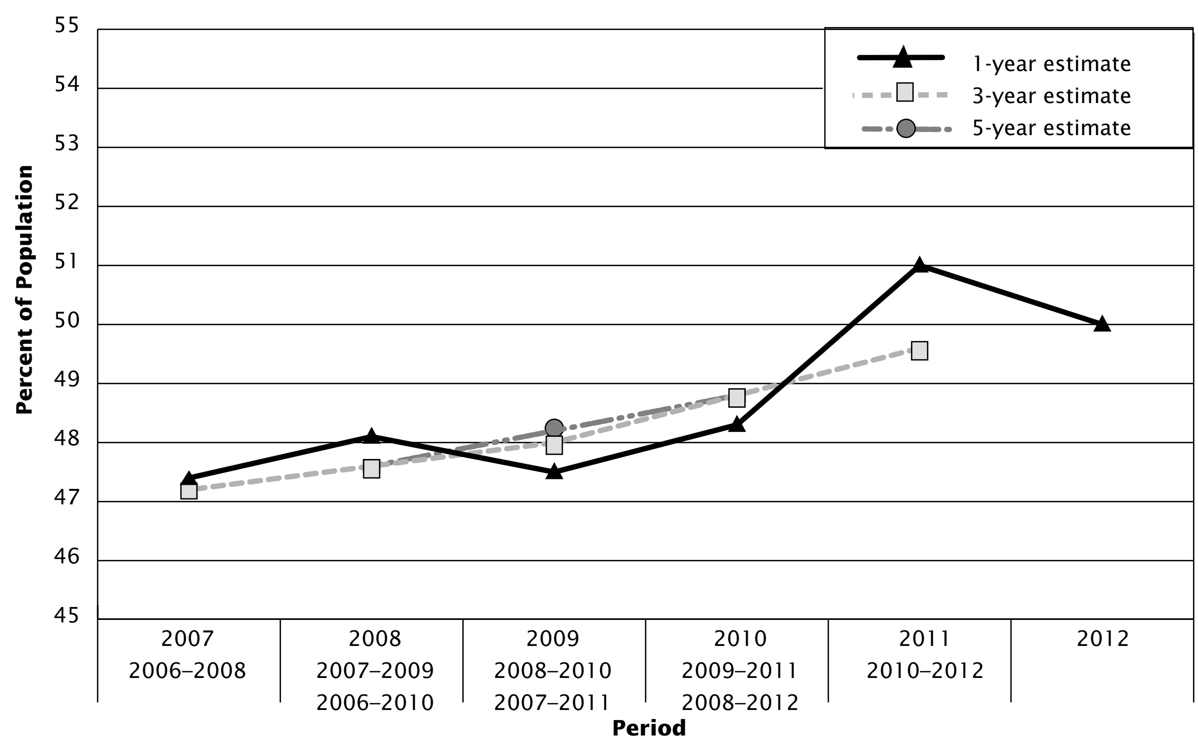
Local area variations may occur due to rapidly occurring changes. As discussed previously, multiyear estimates will tend to be insensitive to such changes when they first occur. Single-year estimates, if associated with sufficiently small CVs, can be very valuable in identifying and studying such phenomena. Graphing trends for such areas using single-year, 3-year, and 5-year estimates can take advantage of the strengths of each set of estimates while using other estimates to compensate for the limitations of each set.

Figure 3 provides an illustration of how the various ACS estimates could be graphed together to better understand local area variations.

The multiyear estimates provide a smoothing of the upward trend and likely provide a better portrayal of the change in proportion over time. Correspondingly, as the data used for single-year estimates will be used in the multiyear estimates, an observed change in the upward direction for consecutive single-year estimates could provide an early indicator of changes in the underlying trend that will be seen when the multiyear estimates encompassing the single years become available.

We hope that you will follow these guidelines to determine when to use single-year versus multiyear estimates, taking into account the intended use and CV associated with the estimate. The Census Bureau encourages you to include the MOE along with the estimate when producing reports, in order to provide the reader with information concerning the uncertainty associated with the estimate.

Figure 3. **Proportion of Population With Bachelor's Degree or Higher, City X Single-Year, Multiyear Estimates**



Source: U.S. Census Bureau. Based on data from the Multiyear Estimates Study.

Appendix 2.

Differences Between ACS and Decennial Census Sample Data

There are many similarities between the methods used in the decennial census sample and the ACS. Both the ACS and the decennial census sample data are based on information from a sample of the population. The data from the Census 2000 sample of about one-sixth of the population were collected using a "long-form" questionnaire, whose content was the model for the ACS. While some differences exist in the specific Census 2000 question wording and that of the ACS, most questions are identical or nearly identical. Differences in the design and implementation of the two surveys are noted below with references provided to a series of evaluation studies that assess the degree to which these differences are likely to impact the estimates. As noted in Appendix 1, the ACS produces period estimates and these estimates do not measure characteristics for the same time frame as the decennial census estimates, which are interpreted to be a snapshot of April 1 of the census year. Additional differences are described below.

Residence Rules, Reference Periods, and Definitions

The fundamentally different purposes of the ACS and the census, and their timing, led to important differences in the choice of data collection methods. For example, the residence rules for a census or survey determine the sample unit's occupancy status and household membership. Defining the rules in a dissimilar way can affect those two very important estimates. The Census 2000 residence rules, which determined where people should be counted, were based on the principle of "usual residence" on April 1, 2000, in keeping with the focus of the census on the requirements of congressional apportionment and state redistricting. To accomplish this the decennial census attempts to restrict and determine a principal place of residence on one specific date for everyone enumerated. The ACS residence rules are based on a "current residence" concept since data are collected continuously throughout the entire year with responses provided relative to the continuously changing survey interview dates. This method is consistent with the goal that the ACS produce estimates that reflect annual averages of the characteristics of all areas.

Estimates produced by the ACS are not measuring exactly what decennial samples have been measuring. The ACS yearly samples, spread over 12 months, collect information that is anchored to the day on which the sampled unit was interviewed, whether it is the day that a mail questionnaire is completed or the day that an interview is conducted by telephone or personal visit. Individual questions with time references such as "last week" or "the last 12 months" all begin the reference period as of this interview date. Even the information on types and amounts of income refers to the 12 months prior to the day the question is answered. ACS interviews are conducted just about every day of the year, and all of the estimates that the survey releases are considered to be averages for a specific time period. The 1-year estimates reflect the full calendar year; 3-year and 5-year estimates reflect the full 36- or 60-month period.

Most decennial census sample estimates are anchored in this same way to the date of enumeration. The most obvious difference between the ACS and the census is the overall time frame in which they are conducted. The census enumeration time period is less than half the time period used to collect data for each single-year ACS estimate. But a more important difference is that the distribution of census enumeration dates are highly clustered in March and April (when most census mail returns were received) with additional, smaller clusters seen in May and June (when nonresponse follow-up activities took place).

This means that the data from the decennial census tend to describe the characteristics of the population and housing in the March through June time period (with an overrepresentation of March/April) while the ACS characteristics describe the characteristics nearly every day over the full calendar year.

Census Bureau analysts have compared sample estimates from Census 2000 with 1-year ACS estimates based on data collected in 2000 and 3-year ACS estimates based on data collected in 1999–2001 in selected counties. A series of reports summarize their findings and can be found at <http://www.census.gov/acs/www/AdvMeth/Reports.htm>. In general, ACS estimates were found to be quite similar to those produced from decennial census data.

More on Residence Rules

Residence rules determine which individuals are considered to be residents of a particular housing unit or group quarters. While many people have definite ties to a single housing unit or group quarters, some people may stay in different places for significant periods of time over the course of the year. For example, migrant workers move with crop seasons and do not live in any one location for the entire year. Differences in treatment of these populations in the census and ACS can lead to differences in estimates of the characteristics of some areas.

For the past several censuses, decennial census residence rules were designed to produce an accurate count of the population as of Census Day, April 1, while the ACS residence rules were designed to collect representative information to produce annual average estimates of the characteristics of all kinds of areas. When interviewing the population living in housing units, the decennial census uses a "usual residence" rule to enumerate people at the place where they live or stay most of the time as of April 1. The ACS uses a "current residence" rule to interview people who are currently living or staying in the sample housing unit as long as their stay at that address will exceed 2 months. The residence rules governing the census enumerations of people in group quarters depend on the type of group quarter and where permitted, whether people claim a "usual residence" elsewhere. The ACS applies a straight de facto residence rule to every type of group quarter. Everyone living or staying in a group quarter on the day it is visited by an ACS interviewer is eligible to be sampled and interviewed for the survey. Further information on residence rules can be found at <http://www.census.gov/acs/www/AdvMeth/CollProc/CollProc1.htm>.

The differences in the ACS and census data as a consequence of the different residence rules are most likely minimal for most areas and most characteristics. However, for certain segments of the population the usual and current residence concepts could result in different residence decisions. Appreciable differences may occur in areas where large proportions of the total population spend several months of the year in what would not be considered their residence under decennial census rules. In particular, data for areas that include large beach, lake, or mountain vacation areas may differ appreciably between the census and the ACS if populations live there for more than 2 months.

More on Reference Periods

The decennial census centers its count and its age distributions on a reference date of April 1, the assumption being that the remaining basic demographic questions also reflect that date, regardless of whether the enumeration is conducted by mail in March or by a field follow-up in July. However, nearly all questions are anchored to the date the interview is provided. Questions with their own reference periods, such as "last week," are referring to the week prior to the interview date. The idea that all census data reflect the characteristics as of April 1 is a myth. Decennial census samples actually provide estimates based on aggregated data reflecting the entire period of decennial data collection, and are greatly influenced by delivery dates of mail questionnaires, success of mail response, and data collection schedules for nonresponse follow-up. The ACS reference periods are, in many ways, similar to those in the census in that they reflect the circumstances on the day the data are collected and the individual reference periods of questions relative to that date. However, the ACS estimates represent the average characteristics over a full year (or sets of years), a different time, and reference period than the census.

Some specific differences in reference periods between the ACS and the decennial census are described below. Users should consider the potential impact these different reference periods could have on distributions when comparing ACS estimates with Census 2000.

Those who are interested in more information about differences in reference periods should refer to the Census Bureau's guidance on comparisons that contrasts for each question the specific reference periods used in Census 2000 with those used in the ACS. See <http://www.census.gov/acs/www/UseData/compACS.htm>.

Income Data

To estimate annual income, the Census 2000 long-form sample used the calendar year prior to Census Day as the reference period, and the ACS uses the 12 months prior to the interview date as the reference period. Thus, while Census 2000 collected income information for calendar year 1999, the ACS collects income information for the 12 months preceding the interview date. The responses are a mixture of 12 reference periods ranging from, in the case of the 2006 ACS single-year estimates, the full calendar year 2005 through November 2006. The ACS income responses for each of these reference periods are individually inflation-adjusted to represent dollar values for the ACS collection year.

School Enrollment

The school enrollment question on the ACS asks if a person had "at any time in the last 3 months attended a school or college." A consistent 3-month reference period is used for all interviews. In contrast, Census 2000 asked if a person had "at any time since February 1 attended a school or college." Since Census 2000 data were collected from mid-March to late-August, the reference period could have been as short as about 6 weeks or as long as 7 months.

Utility Costs

The reference periods for two utility cost questions—gas and electricity—differ between Census 2000 and the ACS. The census asked for annual costs, while the ACS asks for the utility costs in the previous month.

Definitions

Some data items were collected by both the ACS and the Census 2000 long form with slightly different definitions that could affect the comparability of the estimates for these items. One example is annual costs for a mobile home. Census 2000 included installment loan costs in the total annual costs but the ACS does not. In this example, the ACS could be expected to yield smaller estimates than Census 2000.

Implementation

While differences discussed above were a part of the census and survey design objectives, other differences observed between ACS and census results were not by design, but due to nonsampling error—differences related to how well the surveys were conducted. Appendix 6 explains nonsampling error in more detail.

The ACS and the census experience different levels and types of coverage error, different levels and treatment of unit and item nonresponse, and different instances of measurement and processing error. Both Census 2000 and the ACS had similar high levels of survey coverage and low levels of unit nonresponse. Higher levels of unit nonresponse were found in the nonresponse follow-up stage of Census 2000. Higher item nonresponse rates were also found in Census 2000. Please see <http://www.census.gov/acs/www/AdvMeth/Reports.htm> for detailed comparisons of these measures of survey quality.

Appendix 3.
Measures of Sampling Error

All survey and census estimates include some amount of error. Estimates generated from sample survey data have uncertainty associated with them due to their being based on a sample of the population rather than the full population. This uncertainty, referred to as sampling error, means that the estimates derived from a sample survey will likely differ from the values that would have been obtained if the entire population had been included in the survey, as well as from values that would have been obtained had a different set of sample units been selected. All other forms of error are called nonsampling error and are discussed in greater detail in Appendix 6.

Sampling error can be expressed quantitatively in various ways, four of which are presented in this appendix—standard error, margin of error, confidence interval, and coefficient of variation. As the ACS estimates are based on a sample survey of the U.S. population, information about the sampling error associated with the estimates must be taken into account when analyzing individual estimates or comparing pairs of estimates across areas, population subgroups, or time periods. The information in this appendix describes each of these sampling error measures, explaining how they differ and how each should be used. It is intended to assist the user with analysis and interpretation of ACS estimates. Also included are instructions on how to compute margins of error for user-derived estimates.

Sampling Error Measures and Their Derivations

Standard Errors

A standard error (SE) measures the variability of an estimate due to sampling. Estimates derived from a sample (such as estimates from the ACS or the decennial census long form) will generally not equal the population value, as not all members of the population were measured in the survey. The SE provides a quantitative measure of the extent to which an estimate derived from the sample survey can be expected to deviate from this population value. It is the foundational measure from which other sampling error measures are derived. The SE is also used when comparing estimates to determine whether the differences between the estimates can be said to be statistically significant.

A very basic example of the standard error is a population of three units, with values of 1, 2, and 3. The average value for this population is 2. If a simple random sample of size two were selected from this population, the estimates of the average value would be 1.5 (units with values of 1 and 2 selected), 2 (units with values of 1 and 3 selected), or 2.5 (units with values of 2 and 3 selected). In this simple example, two of the three samples yield estimates that do not equal the population value (although the average of the estimates across all possible samples do equal the population value). The standard error would provide an indication of the extent of this variation.

The SE for an estimate depends upon the underlying variability in the population for the characteristic and the sample size used for the survey. In general, the larger the sample size, the smaller the standard error of the estimates produced from the sample. This relationship between sample size and SE is the reason ACS estimates for less populous areas are only published using multiple years of data: to take advantage of the larger sample size that results from aggregating data from more than one year.

Margins of Error

A margin of error (MOE) describes the precision of the estimate at a given level of confidence. The confidence level associated with the MOE indicates the likelihood that the sample estimate is within a certain distance (the MOE) from the population value. Confidence levels of 90 percent, 95 percent, and 99 percent are commonly used in practice to lessen the risk associated with an incorrect inference. The MOE provides a concise measure of the precision of the sample estimate in a table and is easily used to construct confidence intervals and test for statistical significance.

The Census Bureau statistical standard for published data is to use a 90-percent confidence level. Thus, the MOEs published with the ACS estimates correspond to a 90-percent confidence level. However, users may want to use other confidence levels, such as 95 percent or 99 percent. The choice of confidence level is usually a matter of preference, balancing risk for the specific application, as a 90-percent confidence level implies a 10 percent chance of an incorrect inference, in contrast with a 1 percent chance if using a 99-percent confidence level. Thus, if the impact of an incorrect conclusion is substantial, the user should consider increasing the confidence level.

One commonly experienced situation where use of a 95 percent or 99 percent MOE would be preferred is when conducting a number of tests to find differences between sample estimates. For example, if one were conducting comparisons between male and female incomes for each of 100 counties in a state, using a 90-percent confidence level would imply that 10 of the comparisons would be expected to be found significant even if no differences actually existed. Using a 99-percent confidence level would reduce the likelihood of this kind of false inference.

Calculating Margins of Error for Alternative Confidence Levels

If you want to use an MOE corresponding to a confidence level other than 90 percent, the published MOE can easily be converted by multiplying the published MOE by an adjustment factor. If the desired confidence level is 95 percent, then the factor is equal to 1.960/1.645.[1] If the desired confidence level is 99 percent, then the factor is equal to 2.576/1.645.

Conversion of the published ACS MOE to the MOE for a different confidence level can be expressed as

$$MOE_{95} = \frac{1.960}{1.645} MOE_{ACS}$$

$$MOE_{99} = \frac{2.576}{1.645} MOE_{ACS}$$

where MOE_{ACS} is the ACS published 90 percent MOE for the estimate.

Factors Associated With Margins of Error for Commonly Used Confidence Levels
90 Percent: 1.645 95 Percent: 1.960 99 Percent: 2.576
Census Bureau standard for published MOE is 90 percent.

For example, the ACS published MOE for the 2006 ACS estimated number of civilian veterans in the state of Virginia is ±12,357. The MOE corresponding to a 95-percent confidence level would be derived as follows:

$$MOE_{95} = \frac{1.960}{1.645}\left(\pm 12{,}357\right) = \pm 14{,}723$$

Deriving the Standard Error From the MOE

When conducting exact tests of significance (as discussed in Appendix 4) or calculating the CV for an estimate, the SEs of the estimates are needed. To derive the SE, simply divide the positive value of the published MOE by 1.645.[2]

Derivation of SEs can thus be expressed as

$$SE = \frac{MOE_{ACS}}{1.645}$$

where MOE_{ACS} is the positive value of the ACS published MOE for the estimate.

For example, the ACS published MOE for estimated number of civilian veterans in the state of Virginia from the 2006 ACS is ±12,357. The SE for the estimate would be derived as

$$SE = \frac{12{,}357}{1.645} = 7{,}512$$

Confidence Intervals

A confidence interval (CI) is a range that is expected to contain the average value of the characteristic that would result over all possible samples with a known probability. This probability is called the “level of confidence” or “confidence level.” CIs are useful when graphing estimates to display their sampling variabilites. The sample estimate and its MOE are used to construct the CI.

Constructing a Confidence Interval From a Margin of Error

To construct a CI at the 90-percent confidence level, the published MOE is used. The CI boundaries are determined by adding to and subtracting from a sample estimate, the estimate’s MOE.

For example, if an estimate of 20,000 had an MOE at the 90-percent confidence level of ±1,645, the CI would range from 18,355 (20,000 – 1,645) to 21,645 (20,000 + 1,645).

For CIs at the 95-percent or 99-percent confidence level, the appropriate MOE must first be derived as explained previously.

Construction of the lower and upper bounds for the CI can be expressed as

$$L_{CL} = \hat{X} - MOE_{CL}$$

$$U_{CL} = \hat{X} + MOE_{CL}$$

where $\hat{X}$ is the ACS estimate and

MOE_{CL} is the positive value of the MOE for the estimate at the desired confidence level.

The CI can thus be expressed as the range

$$CI_{CL} = \left(L_{CL}, U_{CL}\right).$$ [3]

[1] The value 1.65 must be used for ACS single-year estimates for 2005 or earlier, as that was the value used to derive the published margin of error from the standard error in those years.

[2] If working with ACS 1-year estimates for 2005 or earlier, use the value 1.65 rather than 1.645 in the adjustment factor.

[3] Users are cautioned to consider logical boundaries when creating confidence intervals from the margins of error. For example, a small population estimate may have a calculated lower bound less than zero. A negative number of persons doesn’t make sense, so the lower bound should be set to zero instead.

For example, to construct a CI at the 95-percent confidence level for the number of civilian veterans in the state of Virginia in 2006, one would use the 2006 estimate (771,782) and the corresponding MOE at the 95-percent confidence level derived above (±14,723).

$$L_{95} = 771{,}782 - 14{,}723 = 757{,}059$$
$$U_{95} = 771{,}782 + 14{,}723 = 786{,}505$$

The 95-percent CI can thus be expressed as the range 757,059 to 786,505.

The CI is also useful when graphing estimates, to show the extent of sampling error present in the estimates, and for visually comparing estimates. For example, given the MOE at the 90-percent confidence level used in constructing the CI above, the user could be 90 percent certain that the value for the population was between 18,355 and 21,645. This CI can be represented visually as

Coefficients of Variation

A coefficient of variation (CV) provides a measure of the relative amount of sampling error that is associated with a sample estimate. The CV is calculated as the ratio of the SE for an estimate to the estimate itself and is usually expressed as a percent. It is a useful barometer of the stability, and thus the usability of a sample estimate. It can also help a user decide whether a single-year or multiyear estimate should be used for analysis. The method for obtaining the SE for an estimate was described earlier.

The CV is a function of the overall sample size and the size of the population of interest. In general, as the estimation period increases, the sample size increases and therefore the size of the CV decreases. A small CV indicates that the sampling error is small relative to the estimate, and thus the user can be more confident that the estimate is close to the population value. In some applications a small CV for an estimate is desirable and use of a multiyear estimate will therefore be preferable to the use of a 1-year estimate that doesn't meet this desired level of precision.

For example, if an estimate of 20,000 had an SE of 1,000, then the CV for the estimate would be 5 percent ([1,000 /20,000] x 100). In terms of usability, the estimate is very reliable. If the CV was noticeably larger, the usability of the estimate could be greatly diminished.

While it is true that estimates with high CVs have important limitations, they can still be valuable as building blocks to develop estimates for higher levels of aggregation. Combining estimates across geographic areas or collapsing characteristic detail can improve the reliability of those estimates as evidenced by reductions in the CVs.

Calculating Coefficients of Variation From Standard Errors

The CV can be expressed as

$$CV = \frac{SE}{\hat{X}} \times 100$$

where $\hat{X}$ is the ACS estimate and SE is the derived SE for the ACS estimate.

For example, to determine the CV for the estimated number of civilian veterans in the state of Virginia in 2006, one would use the 2006 estimate (771,782), and the SE derived previously (7,512).

$$CV = \frac{7{,}512}{771{,}782} \times 100 = 0.1\%$$

This means that the amount of sampling error present in the estimate is only one-tenth of 1 percent the size of the estimate.

The text box below summarizes the formulas used when deriving alternative sampling error measures from the margin or error published with ACS estimates.

Deriving Sampling Error Measures From Published MOE

Margin Error (MOE) for Alternate Confidence Levels

$$MOE_{95} = \frac{1.960}{1.645} MOE_{ACS}$$

$$MOE_{99} = \frac{2.576}{1.645} MOE_{ACS}$$

Standard Error (SE)

$$SE = \frac{MOE_{ACS}}{1.645}$$

Confidence Interval (CI)

$$CI_{CL} = \left(\hat{X} - MOE_{CL}, \hat{X} + MOE_{CL}\right)$$

Coefficient of Variation (CV)

$$CV = \frac{SE}{\hat{X}} \times 100$$

Calculating Margins of Error for Derived Estimates

One of the benefits of being familiar with ACS data is the ability to develop unique estimates called derived estimates. These derived estimates are usually based on aggregating estimates across geographic areas or population subgroups for which combined estimates are not published in American FactFinder (AFF) tables (e.g., aggregate estimates for a three-county area or for four age groups not collapsed).

ACS tabulations provided through AFF contain the associated confidence intervals (pre-2005) or margins of error (MOEs) (2005 and later) at the 90-percent confidence level. However, when derived estimates are generated (e.g., aggregated estimates, proportions, or ratios not available in AFF), the user must calculate the MOE for these derived estimates. The MOE helps protect against misinterpreting small or nonexistent differences as meaningful.

MOEs calculated based on information provided in AFF for the components of the derived estimates will be at the 90-percent confidence level. If an MOE with a confidence level other than 90 percent is desired, the user should first calculate the MOE as instructed below and then convert the results to an MOE for the desired confidence level as described earlier in this appendix.

Calculating MOEs for Aggregated Count Data

To calculate the MOE for aggregated count data:

1) Obtain the MOE of each component estimate.
2) Square the MOE of each component estimate.
3) Sum the squared MOEs.
4) Take the square root of the sum of the squared MOEs.

The result is the MOE for the aggregated count. Algebraically, the MOE for the aggregated count is calculated as:

$$MOE_{agg} = \pm\sqrt{\sum_{c} MOE_c^2}$$

where MOE_c is the MOE of the c^{th} component estimate.

The example below shows how to calculate the MOE for the estimated total number of females living alone in the three Virginia counties/independent cities that border Washington, DC (Fairfax and Arlington counties, Alexandria city) from the 2006 ACS.

Table 1. **Data for Example 1**

Characteristic	Estimate	MOE
Females living alone in Fairfax County (Component 1)	52,354	±3,303
Females living alone in Arlington County (Component 2)	19,464	±2,011
Females living alone in Alexandria city (Component 3)	17,190	±1,854

The aggregate estimate is:

$$\hat{X} = \hat{X}_{Fairfax} + \hat{X}_{Arlington} + \hat{X}_{Alexandria} =$$
$$52{,}354 + 19{,}464 + 17{,}190 = 89{,}008$$

Obtain MOEs of the component estimates:

$$MOE_{Fairfax} = \pm 3{,}303,$$
$$MOE_{Arlington} = \pm 2{,}011,$$
$$MOE_{Alexandria} = \pm 1{,}854$$

Calculate the MOE for the aggregate estimated as the square root of the sum of the squared MOEs.

$$MOE_{agg} = \pm\sqrt{(3{,}303)^2 + (2{,}011)^2 + (1{,}854)^2} =$$
$$\pm\sqrt{18{,}391{,}246} = \pm 4{,}289$$

Thus, the derived estimate of the number of females living alone in the three Virginia counties/independent cities that border Washington, DC, is 89,008, and the MOE for the estimate is ±4,289.

Calculating MOEs for Derived Proportions

The numerator of a proportion is a subset of the denominator (e.g., the proportion of single person households that are female). To calculate the MOE for derived proportions, do the following:

1) Obtain the MOE for the numerator and the MOE for the denominator of the proportion.
2) Square the derived proportion.
3) Square the MOE of the numerator.
4) Square the MOE of the denominator.
5) Multiply the squared MOE of the denominator by the squared proportion.
6) Subtract the result of (5) from the squared MOE of the numerator.
7) Take the square root of the result of (6).
8) Divide the result of (7) by the denominator of the proportion.

The result is the MOE for the derived proportion. Algebraically, the MOE for the derived proportion is calculated as:

$$MOE_p = \frac{\pm\sqrt{MOE_{num}^2 - (\hat{p}^2 * MOE_{den}^2)}}{\hat{X}_{den}}$$

where MOE_{num} is the MOE of the numerator.

MOE_{den} is the MOE of the denominator.

$\hat{p} = \frac{\hat{X}_{num}}{\hat{X}_{den}}$ is the derived proportion.

$\hat{X}_{num}$ is the estimate used as the numerator of the derived proportion.

$\hat{X}_{den}$ is the estimate used as the denominator of the derived proportion.

There are rare instances where this formula will fail—the value under the square root will be negative. If that happens, use the formula for derived ratios in the next section which will provide a conservative estimate of the MOE.

The example below shows how to derive the MOE for the estimated proportion of Black females 25 years of age and older in Fairfax County, Virginia, with a graduate degree based on the 2006 ACS.

Table 2. **Data for Example 2**

Characteristic	Estimate	MOE
Black females 25 years and older with a graduate degree (numerator)	4,634	±989
Black females 25 years and older (denominator)	31,713	±601

The estimated proportion is:

$$\hat{p} = \frac{\hat{X}_{gradBF}}{\hat{X}_{BF}} = \frac{4{,}634}{31{,}713} = 0.1461$$

where $\hat{X}_{gradBF}$ is the ACS estimate of Black females 25 years of age and older in Fairfax County with a graduate degree and $\hat{X}_{BF}$ is the ACS estimate of Black females 25 years of age and older in Fairfax County.

Obtain MOEs of the numerator (number of Black females 25 years of age and older in Fairfax County with a graduate degree) and denominator (number of Black females 25 years of age and older in Fairfax County).

$$MOE_{num} = \pm 989,\ MOE_{den} = \pm 601$$

Multiply the squared MOE of the denominator by the squared proportion and subtract the result from the squared MOE of the numerator.

$$MOE_{num}^2 - (\hat{p}^2 * MOE_{den}^2) =$$
$$(989)^2 - [(0.1461)^2 * (601)^2] =$$
$$978{,}121 - 7{,}712.3 = 970{,}408.7$$

Calculate the MOE by dividing the square root of the prior result by the denominator.

$$MOE_p = \frac{\pm\sqrt{970{,}408.7}}{31{,}373} = \frac{\pm 985.1}{31{,}373} = \pm 0.0311$$

Thus, the derived estimate of the proportion of Black females 25 years of age and older with a graduate degree in Fairfax County, Virginia, is 0.1461, and the MOE for the estimate is ±0.0311.

Calculating MOEs for Derived Ratios

The numerator of a ratio is not a subset (e.g., the ratio of females living alone to males living alone). To calculate the MOE for derived ratios:

1) Obtain the MOE for the numerator and the MOE for the denominator of the ratio.
2) Square the derived ratio.
3) Square the MOE of the numerator.
4) Square the MOE of the denominator.
5) Multiply the squared MOE of the denominator by the squared ratio.
6) Add the result of (5) to the squared MOE of the numerator.
7) Take the square root of the result of (6).
8) Divide the result of (7) by the denominator of the ratio.

The result is the MOE for the derived ratio. Algebraically, the MOE for the derived ratio is calculated as:

$$MOE_R = \frac{\pm\sqrt{MOE_{num}^2 + (\hat{R}^2 * MOE_{den}^2)}}{\hat{X}_{den}}$$

where MOE_{num} is the MOE of the numerator.

MOE_{den} is the MOE of the denominator.

$\hat{R} = \frac{\hat{X}_{num}}{\hat{X}_{den}}$ is the derived ratio.

$\hat{X}_{num}$ is the estimate used as the numerator of the derived ratio.

$\hat{X}_{den}$ is the estimate used as the denominator of the derived ratio.

The example below shows how to derive the MOE for the estimated ratio of Black females 25 years of age and older in Fairfax County, Virginia, with a graduate degree to Black males 25 years and older in Fairfax County with a graduate degree, based on the 2006 ACS.

Table 3. **Data for Example 3**

Characteristic	Estimate	MOE
Black females 25 years and older with a graduate degree (numerator)	4,634	±989
Black males 25 years and older with a graduate degree (denominator)	6,440	±1,328

The estimated ratio is:

$$\hat{R} = \frac{\hat{X}_{gradBF}}{\hat{X}_{gradBM}} = \frac{4{,}634}{6{,}440} = 0.7200$$

Obtain MOEs of the numerator (number of Black females 25 years of age and older with a graduate degree in Fairfax County) and denominator (number of Black males 25 years of age and older in Fairfax County with a graduate degree).

$$MOE_{num} = \pm 989,\ MOE_{den} = \pm 1{,}328$$

Multiply the squared MOE of the denominator by the squared proportion and add the result to the squared MOE of the numerator.

$$MOE_{num}^2 + (\hat{R}^2 * MOE_{den}^2) =$$

$$(989)^2 + [(0.7200)^2 * (1{,}328)^2] =$$

$$978{,}121 + 913{,}318.1 = 1{,}891{,}259.1$$

Calculate the MOE by dividing the square root of the prior result by the denominator.

$$MOE_R = \frac{\pm\sqrt{1{,}891{,}259.1}}{6{,}440} = \frac{\pm 1{,}375.2}{6{,}440} = \pm 0.2135$$

Thus, the derived estimate of the ratio of the number of Black females 25 years of age and older in Fairfax County, Virginia, with a graduate degree to the number of Black males 25 years of age and older in Fairfax County, Virginia, with a graduate degree is 0.7200, and the MOE for the estimate is ±0.2135.

Calculating MOEs for the Product of Two Estimates

To calculate the MOE for the product of two estimates, do the following:

1) Obtain the MOEs for the two estimates being multiplied together.
2) Square the estimates and their MOEs.
3) Multiply the first squared estimate by the second estimate's squared MOE.
4) Multiply the second squared estimate by the first estimate's squared MOE.
5) Add the results from (3) and (4).
6) Take the square root of (5).

The result is the MOE for the product. Algebraically, the MOE for the product is calculated as:

$$MOE_{A\times B} = \pm\sqrt{A^2 \times MOE_B^2 + B^2 \times MOE_A^2}$$

where *A* and *B* are the first and second estimates, respectively.

MOE_A is the MOE of the first estimate.

MOE_B is the MOE of the second estimate.

The example below shows how to derive the MOE for the estimated number of Black workers 16 years and over in Fairfax County, Virginia, who used public transportation to commute to work, based on the 2006 ACS.

Table 4. **Data for Example 4**

Characteristic	Estimate	MOE
Black workers 16 years and over (first estimate)	50,624	±2,423
Percent of Black workers 16 years and over who commute by public transportation (second estimate)	13.4%	±2.7%

To apply the method, the proportion (0.134) needs to be used instead of the percent (13.4). The estimated product is 50,624 × 0.134 = 6,784. The MOE is calculated by:

$$MOE_{A\times B} = \pm\sqrt{50{,}624^2 \times 0.027^2 + 0.134^2 \times 2{,}423^2}$$

$$= \pm 1{,}405$$

Thus, the derived estimate of Black workers 16 years and over who commute by public transportation is 6,784, and the MOE of the estimate is ±1,405.

Calculating MOEs for Estimates of "Percent Change" or "Percent Difference"

The "percent change" or "percent difference" between two estimates (for example, the same estimates in two different years) is commonly calculated as

$$Percent\ Change = 100\% * \frac{\hat{X}_2 - \hat{X}_1}{\hat{X}_1}$$

Because $\hat{X}_2$ is not a subset of $\hat{X}_1$, the procedure to calculate the MOE of a ratio discussed previously should be used here to obtain the MOE of the percent change.

The example below shows how to calculate the margin of error of the percent change using the 2006 and 2005 estimates of the number of persons in Maryland who lived in a different house in the U.S. 1 year ago.

Table 5. **Data for Example 5**

Characteristic	Estimate	MOE
Persons who lived in a different house in the U.S. 1 year ago, 2006	802,210	±22,866
Persons who lived in a different house in the U.S. 1 year ago, 2005	762,475	±22,666

The percent change is:

$$Percent\ Change = 100\% * \frac{\hat{X}_2 - \hat{X}_1}{\hat{X}_1} =$$

$$100\% * \left(\frac{802{,}210 - 762{,}475}{762{,}475}\right) = 5.21\%$$

For use in the ratio formula, the ratio of the two estimates is:

$$\hat{R} = \frac{\hat{X}_2}{\hat{X}_1} = \frac{802{,}210}{762{,}475} = 1.0521$$

The MOEs for the numerator ($\hat{X}_2$) and denominator ($\hat{X}_1$) are:

$$MOE_2 = +/-22{,}866,\ MOE_1 = +/-22{,}666$$

Add the squared MOE of the numerator (MOE_2) to the product of the squared ratio and the squared MOE of the denominator (MOE_1):

$$MOE_2^2 + (\hat{R}^2 * MOE_1^2) =$$

$$(22{,}866)^2 + [(1.0521)^2 * (22{,}666)^2] =$$

$$1{,}091{,}528{,}529$$

Calculate the MOE by dividing the square root of the prior result by the denominator ($\hat{X}_1$).

$$MOE_R = \frac{\pm\sqrt{1{,}091{,}528{,}529}}{762{,}475} = \frac{\pm 33{,}038.3}{762{,}475} = \pm 0.0433$$

Finally, the MOE of the percent change is the MOE of the ratio, multiplied by 100 percent, or 4.33 percent.

The text box below summarizes the formulas used to calculate the margin of error for several derived estimates.

Calculating Margins of Error for Derived Estimates

Aggregated Count Data

$$MOE_{agg} = \pm\sqrt{\sum_c MOE_c^2}$$

Derived Proportions

$$MOE_p = \frac{\pm\sqrt{MOE_{num}^2 - (\hat{p}^2 * MOE_{den}^2)}}{\hat{X}_{den}}$$

Derived Ratios

$$MOE_R = \frac{\pm\sqrt{MOE_{num}^2 + (\hat{R}^2 * MOE_{den}^2)}}{\hat{X}_{den}}$$

Appendix 4.
Making Comparisons

One of the most important uses of the ACS estimates is to make comparisons between estimates. Several key types of comparisons are of general interest to users: 1) comparisons of estimates from different geographic areas within the same time period (e.g., comparing the proportion of people below the poverty level in two counties); 2) comparisons of estimates for the same geographic area across time periods (e.g., comparing the proportion of people below the poverty level in a county for 2006 and 2007); and 3) comparisons of ACS estimates with the corresponding estimates from past decennial census samples (e.g., comparing the proportion of people below the poverty level in a county for 2006 and 2000).

A number of conditions must be met when comparing survey estimates. Of primary importance is that the comparison takes into account the sampling error associated with each estimate, thus determining whether the observed differences between estimates are statistically significant. Statistical significance means that there is statistical evidence that a true difference exists within the full population, and that the observed difference is unlikely to have occurred by chance due to sampling. A method for determining statistical significance when making comparisons is presented in the next section. Considerations associated with the various types of comparisons that could be made are also discussed.

Determining Statistical Significance

When comparing two estimates, one should use the test for significance described below. This approach will allow the user to ascertain whether the observed difference is likely due to chance (and thus is not statistically significant) or likely represents a true difference that exists in the population as a whole (and thus is statistically significant).

The test for significance can be carried out by making several computations using the estimates and their corresponding standard errors (SEs). When working with ACS data, these computations are simple given the data provided in tables in the American FactFinder.

1) Determine the SE for each estimate (for ACS data, SE is defined by the positive value of the margin of error (MOE) divided by 1.645).[4]
2) Square the resulting SE for each estimate.
3) Sum the squared SEs.
4) Calculate the square root of the sum of the squared SEs.
5) Calculate the difference between the two estimates.
6) Divide (5) by (4).
7) Compare the absolute value of the result of (6) with the critical value for the desired level of confidence (1.645 for 90 percent, 1.960 for 95 percent, 2.576 for 99 percent).
8) If the absolute value of the result of (6) is greater than the critical value, then the difference between the two estimates can be considered statistically significant at the level of confidence corresponding to the critical value used in (7).

Algebraically, the significance test can be expressed as follows:

If $\left|\frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}}\right| > Z_{CL}$, then the difference

between estimates $\hat{X}_1$ and $\hat{X}_2$ is statistically significant at the specified confidence level, CL

where $\hat{X}_i$ is estimate *i* (=1,2)

SE_i is the SE for the estimate *i* (=1,2)

Z_{CL} is the critical value for the desired confidence level (=1.645 for 90 percent, 1.960 for 95 percent, 2.576 for 99 percent).

The example below shows how to determine if the difference in the estimated percentage of households in 2006 with one or more people of age 65 and older between State A (estimated percentage =22.0, SE=0.12) and State B (estimated percentage =21.5, SE=0.12) is statistically significant. Using the formula above:

$$\left|\frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}}\right| = \left|\frac{22.0 - 21.5}{\sqrt{(0.12)^2 + (0.12)^2}}\right| =$$

$$\left|\frac{0.5}{\sqrt{0.015 + 0.015}}\right| = \left|\frac{0.5}{\sqrt{0.03}}\right| = \left|\frac{0.5}{0.173}\right| = 2.90$$

Since the test value (2.90) is greater than the critical value for a confidence level of 99 percent (2.576), the difference in the percentages is statistically significant at a 99-percent confidence level. This is also referred to as statistically significant at the alpha = 0.01 level. A rough interpretation of the result is that the user can be 99 percent certain that a difference exists between the percentages of households with one or more people aged 65 and older between State A and State B.

[4] NOTE: If working with ACS single-year estimates for 2005 or earlier, use the value 1.65 rather than 1.645.

By contrast, if the corresponding estimates for State C and State D were 22.1 and 22.5, respectively, with standard errors of 0.20 and 0.25, respectively, the formula would yield

$$\left|\frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}}\right| = \left|\frac{22.5 - 22.1}{\sqrt{(0.20)^2 + (0.25)^2}}\right| =$$

$$\left|\frac{0.4}{\sqrt{0.04 + 0.0625}}\right| = \left|\frac{0.4}{\sqrt{0.1025}}\right| = \left|\frac{0.4}{0.320}\right| = 1.25$$

Since the test value (1.25) is less than the critical value for a confidence level of 90 percent (1.645), the difference in percentages is not statistically significant. A rough interpretation of the result is that the user cannot be certain to any sufficient degree that the observed difference in the estimates was not due to chance.

Comparisons Within the Same Time Period

Comparisons involving two estimates from the same time period (e.g., from the same year or the same 3-year period) are straightforward and can be carried out as described in the previous section. There is, however, one statistical aspect related to the test for statistical significance that users should be aware of. When comparing estimates within the same time period, the areas or groups will generally be nonoverlapping (e.g., comparing estimates for two different counties). In this case, the two estimates are independent, and the formula for testing differences is statistically correct.

In some cases, the comparison may involve a large area or group and a subset of the area or group (e.g., comparing an estimate for a state with the corresponding estimate for a county within the state or comparing an estimate for all females with the corresponding estimate for Black females). In these cases, the two estimates are not independent. The estimate for the large area is partially dependent on the estimate for the subset and, strictly speaking, the formula for testing differences should account for this partial dependence. However, unless the user has reason to believe that the two estimates are strongly correlated, it is acceptable to ignore the partial dependence and use the formula for testing differences as provided in the previous section. However, if the two estimates are positively correlated, a finding of statistical significance will still be correct, but a finding of a lack of statistical significance based on the formula may be incorrect. If it is important to obtain a more exact test of significance, the user should consult with a statistician about approaches for accounting for the correlation in performing the statistical test of significance.

Comparisons Across Time Periods

Comparisons of estimates from different time periods may involve different single-year periods or different multiyear periods of the same length within the same area. Comparisons across time periods should be made only with comparable time period estimates. Users are advised against comparing single-year estimates with multiyear estimates (e.g., comparing 2006 with 2007–2009) and against comparing multiyear estimates of differing lengths (e.g., comparing 2006–2008 with 2009–2014), as they are measuring the characteristics of the population in two different ways, so differences between such estimates are difficult to interpret. When carrying out any of these types of comparisons, users should take several other issues into consideration.

When comparing estimates from two different single-year periods, one prior to 2006 and the other 2006 or later (e.g., comparing estimates from 2005 and 2007), the user should recognize that from 2006 on the ACS sample includes the population living in group quarters (GQ) as well as the population living in housing units. Many types of GQ populations have demographic, social, or economic characteristics that are very different from the household population. As a result, comparisons between 2005 and 2006 and later ACS estimates could be affected. This is particularly true for areas with a substantial GQ population. For most population characteristics, the Census Bureau suggests users make comparisons across these time periods only if the geographic area of interest does not include a substantial GQ population. For housing characteristics or characteristics published only for the household population, this is obviously not an issue.

Comparisons Based on Overlapping Periods

When comparing estimates from two multiyear periods, ideally comparisons should be based on nonoverlapping periods (e.g., comparing estimates from 2006–2008 with estimates from 2009–2011). The comparison of two estimates for different, but overlapping periods is challenging since the difference is driven by the nonoverlapping years. For example, when comparing the 2005–2007 ACS with the 2006–2008 ACS, data for 2006 and 2007 are included in both estimates. Their contribution is subtracted out when the estimate of differences is calculated. While the interpretation of this difference is difficult, these comparisons can be made with caution. Under most circumstances, the estimate of difference should not be interpreted as a reflection of change between the last 2 years.

The use of MOEs for assessing the reliability of change over time is complicated when change is being evaluated using multiyear estimates. From a technical standpoint, change over time is best evaluated with multiyear estimates that do not overlap. At the same time, many areas whose only source of data will be 5-year estimates will not want to wait until 2015 to evaluate change (i.e., comparing 2005–2009 with 2010–2014).

When comparing two 3-year estimates or two 5-year estimates of the same geography that overlap in sample years one must account for this sample overlap. Thus to calculate the standard error of this difference use the following approximation to the standard error:

$$SE(\hat{X}_1 - \hat{X}_2) \cong \sqrt{(1-C)}\sqrt{SE_1^{\,2} + SE_2^{\,2}}$$

where C is the fraction of overlapping years. For example, the periods 2005–2009 and 2007–2011 overlap for 3 out of 5 years, so C=3/5=0.6. If the periods do not overlap, such as 2005–2007 and 2008–2010, then C=0.

With this SE one can test for the statistical significance of the difference between the two estimates using the method outlined in the previous section with one modification; substitute $\sqrt{(1-C)}\sqrt{SE_1^{\,2} + SE_2^{\,2}}$ for $\sqrt{SE_1^{\,2} + SE_2^{\,2}}$ in the denominator of the formula for the significance test.

Comparisons With Census 2000 Data

In Appendix 2, major differences between ACS data and decennial census sample data are discussed. Factors such as differences in residence rules, universes, and reference periods, while not discussed in detail in this appendix, should be considered when comparing ACS estimates with decennial census estimates. For example, given the reference period differences, seasonality may affect comparisons between decennial census and ACS estimates when looking at data for areas such as college towns and resort areas.

The Census Bureau subject matter specialists have reviewed the factors that could affect differences between ACS and decennial census estimates and they have determined that ACS estimates are similar to those obtained from past decennial census sample data for most areas and characteristics. The user should consider whether a particular analysis involves an area or characteristic that might be affected by these differences.[5]

When comparing ACS and decennial census sample estimates, the user must remember that the decennial census sample estimates have sampling error associated with them and that the standard errors for both ACS and census estimates must be incorporated when performing tests of statistical significance. Appendix 3 provides the calculations necessary for determining statistical significance of a difference between two estimates. To derive the SEs of census sample estimates, use the method described in Chapter 8 of either the Census 2000 Summary File 3 Technical Documentation <http://www.census.gov/prod/cen2000/doc/sf3.pdf> or the Census 2000 Summary File 4 Technical Documentation <http://www.census.gov/prod/cen2000/doc/sf4.pdf>.

A conservative approach to testing for statistical significance when comparing ACS and Census 2000 estimates that avoids deriving the SE for the Census 2000 estimate would be to assume the SE for the Census 2000 estimate is the same as that determined for the ACS estimate. The result of this approach would be that a finding of statistical significance can be assumed to be accurate (as the SE for the Census 2000 estimate would be expected to be less than that for the ACS estimate), but a finding of no statistical significance could be incorrect. In this case the user should calculate the census long-form standard error and follow the steps to conduct the statistical test.

Comparisons With 2010 Census Data

Looking ahead to the 2010 decennial census, data users need to remember that the socioeconomic data previously collected on the long form during the census will not be available for comparison with ACS estimates. The only common variables for the ACS and 2010 Census are sex, age, race, ethnicity, household relationship, housing tenure, and vacancy status.

The critical factor that must be considered when comparing ACS estimates encompassing 2010 with the 2010 Census is the potential impact of housing and population controls used for the ACS. As the housing and population controls used for 2010 ACS data will be based on the Population Estimates Program where the estimates are benchmarked on the Census 2000 counts, they will not agree with the 2010 Census population counts for that year. The 2010 population estimates may differ from the 2010 Census counts for two major reasons—the true change from 2000 to 2010 is not accurately captured by the estimates and the completeness of coverage in the 2010 Census is different than coverage of Census 2000. The impact of this difference will likely affect most areas and states, and be most notable for smaller geographic areas where the potential for large differences between the population controls and the 2010 Census population counts is greater.

Comparisons With Other Surveys

Comparisons of ACS estimates with estimates from other national surveys, such as the Current Population Survey, may be of interest to some users. A major consideration in making such comparisons will be that ACS

[5] Further information concerning areas and characteristics that do not fit the general pattern of comparability can be found on the ACS Web site at <http://www.census.gov/acs/www/UseData/compACS.htm>.

estimates include data for populations in both institutional and noninstitutional group quarters, and estimates from most national surveys do not include institutional populations. Another potential for large effects when comparing data from the ACS with data from other national surveys is the use of different questions for measuring the same or similar information.

Sampling error and its impact on the estimates from the other survey should be considered if comparisons and statements of statistical difference are to be made, as described in Appendix 3. The standard errors on estimates from other surveys should be derived according to technical documentation provided for those individual surveys.

Finally, the user wishing to compare ACS estimates with estimates from other national surveys should consider the potential impact of other factors, such as target population, sample design and size, survey period, reference period, residence rules, and interview modes on estimates from the two sources.

Appendix 5.
Using Dollar-Denominated Data

Dollar-denominated data refer to any characteristics for which inflation adjustments are used when producing annual estimates. For example, income, rent, home value, and energy costs are all dollar-denominated data.

Inflation will affect the comparability of dollar-denominated data across time periods. When ACS multiyear estimates for dollar-denominated data are generated, amounts are adjusted using inflation factors based on the Consumer Price Index (CPI).

Given the potential impact of inflation on observed differences of dollar-denominated data across time periods, users should adjust for the effects of inflation. Such an adjustment will provide comparable estimates accounting for inflation. In making adjustments, the Census Bureau recommends using factors based on the All Items CPI-U-RS (CPI research series). The Bureau of Labor Statistics CPI indexes through 2006 are found at <http://www.bls.gov/cpi/cpiurs1978_2006.pdf>. Explanations follow.

Creating Single-Year Income Values

ACS income values are reported based on the amount of income received during the 12 months preceding the interview month. This is the income reference period. Since there are 12 different income reference periods throughout an interview year, 12 different income inflation adjustments are made. Monthly CPI-U-RSs are used to inflation-adjust the 12 reference period incomes to a single reference period of January through December of the interview year. Note that there are no inflation adjustments for single-year estimates of rent, home value, or energy cost values.

Adjusting Single-Year Estimates Over Time

When comparing single-year income, rent, home value, and energy cost value estimates from two different years, adjustment should be made as follows:

1) Obtain the All Items CPI-U-RS Annual Averages for the 2 years being compared.

2) Calculate the inflation adjustment factor as the ratio of the CPI-U-RS from the more recent year to the CPI-U-RS from the earlier year.

3) Multiply the dollar-denominated data estimated for the earlier year by the inflation adjustment factor.

The inflation-adjusted estimate for the earlier year can be expressed as:

$$\hat{X}_{Y1,Adj} = \frac{CPI_{Y2}}{CPI_{Y1}} \hat{X}_{Y1}$$

where CPI_{Y1} is the All Items CPI-U-RS Annual Average for the earlier year (Y1).

CPI_{Y2} is the All Items CPI-U-RS Annual Average for the more recent year (Y2).

$\hat{X}_{Y1}$ is the published ACS estimate for the earlier year (Y1).

The example below compares the national median value for owner-occupied mobile homes in 2005 ($37,700) and 2006 ($41,000). First adjust the 2005 median value using the 2005 All Items CPI-U-RS Annual Average (286.7) and the 2006 All Items CPI-U-RS Annual Average (296.1) as follows:

$$\hat{X}_{2005,Adj} = \frac{296.1}{286.7} \times \$37{,}700 = \$38{,}936$$

Thus, the comparison of the national median value for owner-occupied mobile homes in 2005 and 2006, in 2006 dollars, would be $38,936 (2005 inflation-adjusted to 2006 dollars) versus $41,000 (2006 dollars).

Creating Values Used in Multiyear Estimates

Multiyear income, rent, home value, and energy cost values are created with inflation adjustments. The Census Bureau uses the All Items CPI-U-RS Annual Averages for each year in the multiyear time period to calculate a set of inflation adjustment factors. Adjustment factors for a time period are calculated as ratios of the CPI-U-RS Annual Average from its most recent year to the CPI-U-RS Annual Averages from each of its earlier years. The ACS values for each of the earlier years in the multiyear period are multiplied by the appropriate inflation adjustment factors to produce the inflation-adjusted values. These values are then used to create the multiyear estimates.

As an illustration, consider the time period 2004–2006, which consisted of individual reference-year income values of $30,000 for 2006, $20,000 for 2005, and $10,000 for 2004. The multiyear income components are created from inflation-adjusted reference period income values using factors based on the All Items CPI-U-RS Annual Averages of 277.4 (for 2004), 286.7 (for 2005), and 296.1 (for 2006). The adjusted 2005 value is the ratio of 296.1 to 286.7 applied to $20,000, which equals $20,656. Similarly, the 2004 value is the ratio of 296.1 to 277.4 applied to $10,000, which equals $10,674.

Adjusting Multiyear Estimates Over Time

When comparing multiyear estimates from two different time periods, adjustments should be made as follows:

1) Obtain the All Items CPI-U-RS Annual Average for the most current year in each of the time periods being compared.
2) Calculate the inflation adjustment factor as the ratio of the CPI-U-RS Annual Average in (1) from the most recent year to the CPI-U-RS in (1) from the earlier years.
3) Multiply the dollar-denominated estimate for the earlier time period by the inflation adjustment factor.

The inflation-adjusted estimate for the earlier years can be expressed as:

$$\hat{X}_{P1,Adj} = \frac{CPI_{P2}}{CPI_{P1}} \hat{X}_{P1}$$

where CPI_{P1} is the All Items CPI-U-RS Annual Average for the last year in the earlier time period (P1).

CPI_{P2} is the All Items CPI-U-RS Annual Average for the last year in the most recent time period (P2).

$\hat{X}_{P1}$ is the published ACS estimate for the earlier time period (P1).

As an illustration, consider ACS multiyear estimates for the two time periods of 2001–2003 and 2004–2006. To compare the national median value for owner-occupied mobile homes in 2001–2003 ($32,000) and 2004–2006 ($39,000), first adjust the 2001–2003 median value using the 2003 All Items CPI-U-RS Annual Averages (270.1) and the 2006 All Items CPI-U-RS Annual Averages (296.1) as follows:

$$\hat{X}_{2001-2003,Adj} = \frac{296.1}{270.1} \times \$32{,}000 = \$35{,}080$$

Thus, the comparison of the national median value for owner-occupied mobile homes in 2001–2003 and 2004–2006, in 2006 dollars, would be $35,080 (2001–2003 inflation-adjusted to 2006 dollars) versus $39,000 (2004–2006, already in 2006 dollars).

Issues Associated With Inflation Adjustment

The recommended inflation adjustment uses a national level CPI and thus will not reflect inflation differences that may exist across geographies. In addition, since the inflation adjustment uses the All Items CPI, it will not reflect differences that may exist across characteristics such as energy and housing costs.

Appendix 6.

Measures of Nonsampling Error

All survey estimates are subject to both sampling and nonsampling error. In Appendix 3, the topic of sampling error and the various measures available for understanding the uncertainty in the estimates due to their being derived from a sample, rather than from an entire population, are discussed. The margins of error published with ACS estimates measure only the effect of sampling error. Other errors that affect the overall accuracy of the survey estimates may occur in the course of collecting and processing the ACS, and are referred to collectively as nonsampling errors.

Broadly speaking, nonsampling error refers to any error affecting a survey estimate outside of sampling error. Nonsampling error can occur in complete censuses as well as in sample surveys, and is commonly recognized as including coverage error, unit nonresponse, item nonresponse, response error, and processing error.

Types of Nonsampling Errors

Coverage error occurs when a housing unit or person does not have a chance of selection in the sample (undercoverage), or when a housing unit or person has more than one chance of selection in the sample, or is included in the sample when they should not have been (overcoverage). For example, if the frame used for the ACS did not allow the selection of newly constructed housing units, the estimates would suffer from errors due to housing undercoverage.

The final ACS estimates are adjusted for under- and overcoverage by controlling county-level estimates to independent total housing unit controls and to independent population controls by sex, age, race, and Hispanic origin (more information is provided on the coverage error definition page of the "ACS Quality Measures" Web site at <http://www.census.gov/acs/www/UseData/sse/cov/cov_def.htm>). However, it is important to measure the extent of coverage adjustment by comparing the precontrolled ACS estimates to the final controlled estimates. If the extent of coverage adjustments is large, there is a greater chance that differences in characteristics of undercovered or overcovered housing units or individuals differ from those eligible to be selected. When this occurs, the ACS may not provide an accurate picture of the population prior to the coverage adjustment, and the population controls may not eliminate or minimize that coverage error.

Unit nonresponse is the failure to obtain the minimum required information from a housing unit or a resident of a group quarter in order for it to be considered a completed interview. Unit nonresponse means that no survey data are available for a particular sampled unit or person. For example, if no one in a sampled housing unit is available to be interviewed during the time frame for data collection, unit nonresponse will result.

It is important to measure unit nonresponse because it has a direct effect on the quality of the data. If the unit nonresponse rate is high, it increases the chance that the final survey estimates may contain bias, even though the ACS estimation methodology includes a nonresponse adjustment intended to control potential unit nonresponse bias. This will happen if the characteristics of nonresponding units differ from the characteristics of responding units.

Item nonresponse occurs when a respondent fails to provide an answer to a required question or when the answer given is inconsistent with other information. With item nonresponse, while some responses to the survey questionnaire for the unit are provided, responses to other questions are not obtained. For example, a respondent may be unwilling to respond to a question about income, resulting in item nonresponse for that question. Another reason for item nonresponse may be a lack of understanding of a particular question by a respondent.

Information on item nonresponse allows users to judge the completeness of the data on which the survey estimates are based. Final estimates can be adversely impacted when item nonresponse is high, because bias can be introduced if the actual characteristics of the people who do not respond to a question differ from those of people who do respond to it. The ACS estimation methodology includes imputations for item nonresponse, intended to reduce the potential for item nonresponse bias.

Response error occurs when data are reported or recorded incorrectly. Response errors may be due to the respondent, the interviewer, the questionnaire, or the survey process itself. For example, if an interviewer conducting a telephone interview incorrectly records a respondent's answer, response error results. In the same way, if the respondent fails to provide a correct response to a question, response error results. Another potential source of response error is a survey process that allows proxy responses to be obtained, wherein a knowledgeable person within the household provides responses for another person within the household who is unavailable for the interview. Even more error prone is allowing neighbors to respond.

Processing error can occur during the preparation of the final data files. For example, errors may occur if data entry of questionnaire information is incomplete or inaccurate. Coding of responses incorrectly also results in processing error. Critical reviews of edits and tabulations by subject matter experts are conducted to keep errors of this kind to a minimum.

Nonsampling error can result in random errors and systematic errors. Of greatest concern are systematic errors. Random errors are less critical since they tend to cancel out at higher geographic levels in large samples such as the ACS.

On the other hand, systematic errors tend to accumulate over the entire sample. For example, if there is an error in the questionnaire design that negatively affects the accurate capture of respondents' answers, processing errors are created. Systematic errors often lead to a bias in the final results. Unlike sampling error and random error resulting from nonsampling error, bias caused by systematic errors cannot be reduced by increasing the sample size.

ACS Quality Measures

Nonsampling error is extremely difficult, if not impossible, to measure directly. However, the Census Bureau has developed a number of indirect measures of nonsampling error to help inform users of the quality of the ACS estimates: sample size, coverage rates, unit response rates and nonresponse rates by reason, and item allocation rates. Starting with the 2007 ACS, these measures are available in the B98 series of detailed tables on AFF. Quality measures for previous years are available on the "ACS Quality Measures" Web site at <http:/www.census.gov/acs/www/UseData/sse/>.

Sample size measures for the ACS summarize information for the housing unit and GQ samples. The measures available at the state level are:[6]

- Housing units
  - Number of initial addresses selected
  - Number of final survey interviews
- Group quarters people (beginning with the 2006 ACS)
  - Number of initial persons selected
  - Number of final survey interviews

Sample size measures may be useful in special circumstances when determining whether to use single-year or multiyear estimates in conjunction with estimates of the population of interest. While the coefficient of variation (CV) should typically be used to determine usability, as explained in Appendix 3, there may be some situations where the CV is small but the user has reason to believe the sample size for a subgroup is very small and the robustness of the estimate is in question.

For example, the Asian-alone population makes up roughly 1 percent (8,418/656,700) of the population in Jefferson County, Alabama. Given that the number of successful housing unit interviews in Jefferson County for the 2006 ACS were 4,072 and assuming roughly 2.5 persons per household (or roughly 12,500 completed person interviews), one could estimate that the 2006 ACS data for Asians in Jefferson County are based on roughly 150 completed person interviews.

Coverage rates are available for housing units, and total population by sex at both the state and national level. Coverage rates for total population by six race/ethnicity categories and the GQ population are also available at the national level. These coverage rates are a measure of the extent of adjustment to the survey weights required during the component of the estimation methodology that adjusts to population controls. Low coverage rates are an indication of greater potential for coverage error in the estimates.

Unit response and nonresponse rates for housing units are available at the county, state, and national level by reason for nonresponse: refusal, unable to locate, no one home, temporarily absent, language problem, other, and data insufficient to be considered an interview. Rates are also provided separately for persons in group quarters at the national and state levels.

A low unit response rate is an indication that there is potential for bias in the survey estimates. For example, the 2006 housing unit response rates are at least 94 percent for all states. The response rate for the District of Columbia in 2006 was 91 percent.

Item allocation rates are determined by the content edits performed on the individual raw responses and closely correspond to item nonresponse rates. Overall housing unit and person characteristic allocation rates are available at the state and national levels, which combine many different characteristics. Allocation rates for individual items may be calculated from the B99 series of imputation detailed tables available in AFF.

Item allocation rates do vary by state, so users are advised to examine the allocation rates for characteristics of interest before drawing conclusions from the published estimates.

[6] The sample size measures for housing units (number of initial addresses selected and number of final survey interviews) and for group quarters people cannot be used to calculate response rates. For the housing unit sample, the number of initial addresses selected includes addresses that were determined not to identify housing units, as well as initial addresses that are subsequently subsampled out in preparation for personal visit nonresponse follow-up. Similarly, the initial sample of people in group quarters represents the expected sample size within selected group quarters prior to visiting and sampling of residents.

Appendix 7.

Implications of Population Controls on ACS Estimates

As with most household surveys, the American Community Survey data are controlled so that the numbers of housing units and people in categories defined by age, sex, race, and Hispanic origin agree with the Census Bureau's official estimates. The American Community Survey (ACS) measures the characteristics of the population, but the official count of the population comes from the previous census, updated by the Population Estimates Program.

In the case of the ACS, the total housing unit estimates and the total population estimates by age, sex, race and Hispanic origin are controlled at the county (or groups of counties) level. The group quarters total population is controlled at the state level by major type of group quarters. Such adjustments are important to correct the survey data for nonsampling and sampling errors. An important source of nonsampling error is the potential under-representation of hard-to-enumerate demographic groups. The use of the population controls results in ACS estimates that more closely reflect the level of coverage achieved for those groups in the preceding census. The use of the population estimates as controls partially corrects demographically implausible results from the ACS due to the ACS data being based on a sample of the population rather than a full count. For example, the use of the population controls "smooths out" demographic irregularities in the age structure of the population that result from random sampling variability in the ACS.

When the controls are applied to a group of counties rather than a single county, the ACS estimates and the official population estimates for the individual counties may not agree. There also may not be agreement between the ACS estimates and the population estimates for levels of geography such as subcounty areas where the population controls are not applied.

The use of population and housing unit controls also reduces random variability in the estimates from year to year. Without the controls, the sampling variability in the ACS could cause the population estimates to increase in one year and decrease in the next (especially for smaller areas or demographic groups), when the underlying trend is more stable. This reduction in variability on a time series basis is important since results from the ACS may be used to monitor trends over time. As more current data become available, the time series of estimates from the Population Estimates Program are revised back to the preceding census while the ACS estimates in previous years are not. Therefore, some differences in the ACS estimates across time may be due to changes in the population estimates.

For single-year ACS estimates, the population and total housing unit estimates for July 1 of the survey year are used as controls. For multiyear ACS estimates, the controls are the average of the individual year population estimates.

Appendix 8.
Other ACS Resources

Background and Overview Information

American Community Survey Web Page Site Map: <http://www.census.gov/acs/www/Site_Map.html> This link is the site map for the ACS Web page. It provides an overview of the links and materials that are available online, including numerous reference documents.

What Is the ACS? <http://www.census.gov/acs/www/SBasics/What/What1.htm> This Web page includes basic information about the ACS and has links to additional information including background materials.

ACS Design, Methodology, Operations

American Community Survey Design and Methodology Technical Paper: <http://www.census.gov/acs/www/Downloads/tp67.pdf> This document describes the basic design of the 2005 ACS and details the full set of methods and procedures that were used in 2005. Please watch our Web site as a revised version will be released in the fall of 2008, detailing methods and procedures used in 2006 and 2007.

About the Data (Methodology: <http://www.census.gov/acs/www/AdvMeth/> This Web page contains links to information on ACS data collection and processing, evaluation reports, multiyear estimates study, and related topics.

ACS Quality

Accuracy of the Data (2007): <http://www.census.gov/acs/www/Downloads/ACS/accuracy2007.pdf> This document provides data users with a basic understanding of the sample design, estimation methodology, and accuracy of the 2007 ACS data.

ACS Sample Size: <http://www.census.gov/acs/www/SBasics/SSizes/SSizes06.htm> This link provides sample size information for the counties that were published in the 2006 ACS. The initial sample size and the final completed interviews are provided. The sample sizes for all published counties and county equivalents starting with the 2007 ACS will only be available in the B98 series of detailed tables on American FactFinder.

ACS Quality Measures: <http://www.census.gov/acs/www/UseData/sse/> This Web page includes information about the steps taken by the Census Bureau to improve the accuracy of ACS data. Four indicators of survey quality are described and measures are provided at the national and state level.

Guidance on Data Products and Using the Data

How to Use the Data: <http://www.census.gov/acs/www/UseData/> This Web page includes links to many documents and materials that explain the ACS data products.

Comparing ACS Data to other sources: <http://www.census.gov/acs/www/UseData/compACS.htm> Tables are provided with guidance on comparing the 2007 ACS data products to 2006 ACS data and Census 2000 data.

Fact Sheet on Using Different Sources of Data for Income and Poverty: <http://www.census.gov/hhes/www/income/factsheet.html> This fact sheet highlights the sources that should be used for data on income and poverty, focusing on comparing the ACS and the Current Population Survey (CPS).

Public Use Microdata Sample (PUMS): <http://www.census.gov/acs/www/Products/PUMS/> This Web page provides guidance in accessing ACS microdata.