

A Compass for Understanding and Using American Community Survey Data

What High School Teachers Need to Know

Issued
December 2008



U S C E N S U S B U R E A U

Helping You Make Informed Decisions

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

United States[®]
Census
2010

Acknowledgments

William P. O'Hare, Senior Fellow, The Annie E. Casey Foundation, **Linda A. Jacobsen**, Vice President, Domestic Programs, Population Reference Bureau, and **Mark Mather**, Associate Vice President, Domestic Programs, Population Reference Bureau, 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., and **Susan P. Love**, Consultant, 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 D. Flannery**, **Cynthia Davis Hollingsworth**, **Susan L. Hostetter**, **Pamela D. Klein**, **Anna M. Owens**, **Clive R. Richmond**, **Enid Santana**, 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.

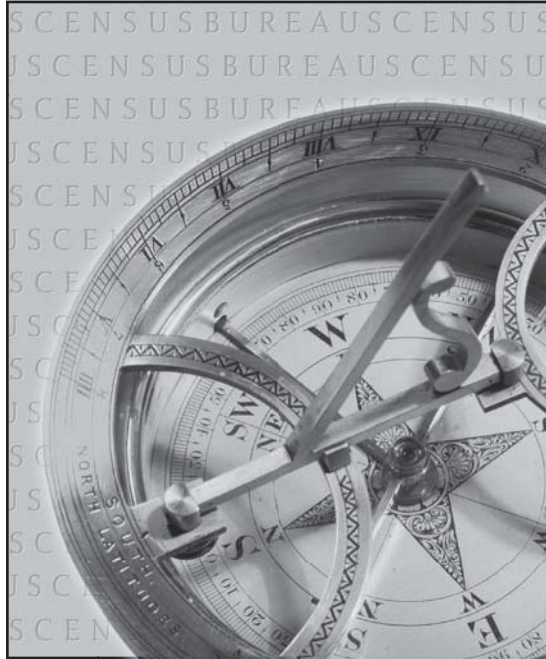
Other individuals who contributed to the review and release of these handbooks include **Dee Alexander**, **Herman Alvarado**, **Mark Asiala**, **Frank Ambrose**, **Maryam Asi**, **Arthur Bakis**, **Genora Barber**, **Michael Beaghen**, **Judy Belton**, **Lisa Blumerman**, **Scott Boggess**, **Ellen Jean Bradley**, **Stephen Buckner**, **Whittona Burrell**, **Edward Castro**, **Gary Chappell**, **Michael Cook**, **Russ Davis**, **Carrie Dennis**, **Jason Devine**, **Joanne Dickinson**, **Barbara Downs**, **Maurice Eleby**, **Sirius Fuller**, **Dale Garrett**, **Yvonne Gist**, **Marjorie Hanson**, **Greg Harper**, **William Hazard**, **Steve Hefter**, **Douglas Hillmer**, **Frank Hobbs**, **Todd Hughes**, **Trina Jenkins**, **Nicholas Jones**, **Anika Juhn**, **Donald Keathley**, **Wayne Kei**, **Karen King**, **Debra 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**, **Dorothy Paugh**, **Marie Pees**, **Marc Perry**, **Greg Pewett**, **Roberto Ramirez**, **Dameka Reese**, **Katherine Reeves**, **Lil Paul Reyes**, **Patrick Rottas**, **Merarys Rios**, **J. Gregory Robinson**, **Anne Ross**, **Marilyn Sanders**, **Nicole Scanniello**, **David Sheppard**, **Joanna Stancil**, **Michael Starsinic**, **Lynette Swopes**, **Anthony Tersine**, **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 Cavis**, Chief, Publications Services Branch, provided general direction and production management.

A Compass for Understanding and Using American Community Survey Data

Issued December 2008

What High School Teachers Need 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 High School Teachers
Need to Know*
U.S. Government Printing Office,
Washington, DC,
2008.



Economics and Statistics Administration

Cynthia A. Glassman,
Under Secretary for Economic Affairs



U.S. CENSUS BUREAU

Steve H. Murdock,
Director

Thomas L. Mesenbourg,
Deputy Director and
Chief Operating Officer

Arnold A. Jackson
Associate Director for Decennial Census

Daniel H. Weinberg
Assistant Director for ACS and Decennial Census

Susan Schechter
Chief, American Community Survey Office

Contents

| | |
|---|------|
| Foreword | iv |
| Introduction | 1 |
| What Is the American Community Survey? | 1 |
| Why the ACS Is Important to High School Teachers and Students | 2 |
| What Information Does the ACS Provide? | 2 |
| How the ACS Works | 3 |
| What Geographic Areas Are Available in the ACS? | 4 |
| Accessing ACS Data | 6 |
| Understanding and Interpreting ACS Data | 12 |
| What Is a Period Estimate? | 12 |
| Interpreting and Comparing Single-Year and Multiyear Estimates | 12 |
| Accuracy of ACS Data | 12 |
| How Teachers Can Use ACS Data | 13 |
| Using the ACS in Social Studies Courses | 13 |
| Using the ACS in Geography Courses | 17 |
| Using the ACS in Mathematics and Statistics Courses | 21 |
| Proportions and Ratios | 22 |
| Statistical Inference | 22 |
| Other Resources for Working With ACS Data | 27 |
| Conclusion | 27 |
| Glossary | 28 |
| Appendixes | A-1 |
| Appendix 1. Understanding and Using Single-Year and Multiyear Estimates | A-1 |
| Appendix 2. Differences Between ACS and Decennial Census Sample Data | A-8 |
| Appendix 3. Measures of Sampling Error..... | A-11 |
| Appendix 4. Making Comparisons | A-18 |
| Appendix 5. Using Dollar-Denominated Data | A-22 |
| Appendix 6. Measures of Nonsampling Error | A-24 |
| Appendix 7. Implications of Population Controls on ACS Estimates..... | A-26 |
| Appendix 8. Other ACS Resources | A-27 |

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

Introduction

Are you a high school teacher looking for new sources of timely information and ways to make your courses more engaging and relevant to students? If so, this handbook will introduce you to a rich, new source of online data: The American Community Survey (ACS). The ACS provides a wide array of social, economic, and demographic information about the nation, your state, and your local communities. These data can be used to teach concepts and skills, such as statistical literacy, and content areas, including social studies, geography, and mathematics. Because the ACS is updated annually, it provides fresh, timely data for your students every year.

This handbook begins with a brief overview of the ACS: what it is, how it came about, and why it is important

to high school teachers and students. We describe the types of information and geographic areas covered by the ACS and explain how to understand and correctly interpret ACS data. Next, we review how to access ACS data online and provide specific examples that illustrate how you can incorporate ACS data into your activities or lesson plans to address a variety of social studies, geography, and mathematics standards. At the end of the handbook, a glossary and a series of technical appendixes are provided that discuss advanced applications and issues with the ACS. Throughout the handbook, we will point readers to these appendixes or other online resources that provide additional or more detailed information about particular topics.

What Is the American Community Survey (ACS)?

As a teacher, you are probably already familiar with the decennial census, conducted every 10 years in the United States since 1790 (see text box). The census provides vital data our country needs to perform governmental functions and to understand the changing social fabric of our nation. Recent censuses have used a “short form” to collect basic information from the entire population and a “long form” to collect detailed

socioeconomic and housing characteristics from a sample of the population (about 1 in every 6 households). Please visit the Census Bureau’s Census in Schools Web site at <http://www.census.gov/dmd/www/teachers.html> where teachers will find lesson plans about the decennial census and free materials for use in their classrooms.

The Importance of the Decennial Census

The decennial census plays a fundamental role in our democracy. The U.S. Constitution stipulates that political power should be based on population size and that the federal government should conduct a census of the entire population every 10 years to apportion seats in the U.S. House of Representatives. This puts the decennial census at the heart of our political process. In the 1960s, the courts determined that decennial census data should also be used to draw both federal congressional and state legislative districts that are equal in population size to make sure every person’s vote carries the same weight (i.e., the “one person-one vote” decision).

For more information about the history of the decennial census, see *Measuring America: The Decennial Censuses from 1790 to 2000*, available online at <http://www.census.gov/prod/www/abs/ma.html>. For more information about how census data are used for redistricting, see *Strength in Numbers: Your Guide to Census 2000 Redistricting Data From the U.S. Census Bureau* available online at http://www.census.gov/rdo/pdf/Guidelines_Censu2000.pdf.

In addition to apportionment and redistricting, decennial census data have also been used for a variety of important functions: government agencies use the information to monitor program effectiveness and distribute federal funds; teachers, journalists, and nonprofit leaders use the data to promote a better understanding of our society; social scientists use the data to conduct research; and businesses use the information to learn more about potential consumers for their products.

Because the pace of change has accelerated so rapidly in the United States, it is no longer sufficient to get detailed data about our population once every 10 years. That is why our government recently started a new population and housing survey called the ACS. The ACS is a continuous, nationwide survey designed to provide timely demographic, housing, and socioeconomic data for states and local communities every year.

The ACS will replace the decennial census long form in 2010 and thereafter by collecting detailed socioeconomic and housing information throughout the decade rather than only once every 10 years.¹ Future decennial censuses will consist of a short form only that will continue to be used to collect the basic data needed for apportionment and redistricting.

Why the ACS Is Important to High School Teachers and Students

Since today's students will become major users of ACS data over the next few decades, it is vital that they have a good understanding of how this new resource can be used, including its strengths and weaknesses. While the ACS provides much of the same data that the decennial census formerly provided, there are some important differences that need to be understood if the new data are to be used effectively. Exposing students to the ACS during high school will help them appreciate the power—and limitations—of this new data source, and how it might serve them in postsecondary education and in their careers. Using the ACS in

the high school classroom may also give students an early exposure to statistical ideas and procedures in a gradual and less intimidating way.

The ACS can bring learning to life for students by giving them up-to-date information about their communities. Like most Americans, students are more interested in the forces that shape their local situation than they are in more abstract and/or national trends and patterns. The ACS can be used to examine a variety of social issues at the state and local levels with timely data.

What Information Does the ACS Provide?

Beginning in 2006, ACS data describe the characteristics of the population living in both housing units and group quarters facilities such as prisons, nursing homes, and college dormitories. Prior to the 2006 ACS, only the housing unit population was surveyed. The ACS covers a wide range of subjects that can help students better understand the social, economic, housing, and demographic characteristics of their communities. The ACS questionnaire covers all of the topics listed in Table 1.

Today's students are accustomed to hands-on, interactive learning. By providing up-to-date information on a wide range of subjects through user-friendly online products, the ACS provides ample opportunities for students to gain first-hand knowledge about important socioeconomic trends in their communities. Whether

the course is focused on current events, geography, sociology, mathematics, economics, or political science, the ACS can help put important national issues into a state and local context. For example, students can compare the racial/ethnic composition of their school district with that of the state or country as a whole, compare median earnings for women and men working full-time in their communities, or use education and earnings statistics to learn about the economic returns of a college degree. ACS data can also be used to teach mathematical and statistical concepts such as proportions, ratios, margins of error, confidence intervals, and coefficients of variation. More options for applying the ACS in the classroom are described later in this handbook in the section titled, "How Teachers Can Use ACS Data."

¹ See Appendix 8 for Web site links to additional background information about the ACS.

Table 1. Subjects Included in the American Community Survey

| Demographic Characteristics | Social Characteristics | Housing Characteristics |
|--|---|----------------------------------|
| Age | Marital Status and Marital History* | Year Structure Built |
| Sex | Fertility | Units in Structure |
| Hispanic Origin | Grandparents as Caregivers | Year Moved Into Unit |
| Race | Ancestry | Rooms |
| Relationship to Householder (e.g., spouse) | Place of Birth, Citizenship, and Year of Entry | Bedrooms |
| | Language Spoken at Home | Kitchen Facilities |
| Economic Characteristics | Educational Attainment and School Enrollment | Plumbing Facilities |
| Income | Residence One Year Ago | House Heating Fuel |
| Food Stamps Benefit | Veteran Status, Period of Military Service, and VA Service-Connected Disability Rating* | Telephone Service Available |
| Labor Force Status | Disability | Farm Residence |
| Industry, Occupation, and Class of Worker | | Financial Characteristics |
| Place of Work and Journey to Work | | Tenure (Owner/Renter) |
| Work Status Last Year | | Housing Value |
| Vehicles Available | | Rent |
| Health Insurance Coverage* | | Selected Monthly Owner Costs |

*Marital History, VA Service-Connected Disability Rating, and Health Insurance Coverage are new for 2008.
Source: U.S. Census Bureau.

How the ACS Works

In 2005, the ACS began sampling nearly 3 million addresses each year, resulting in about 2 million completed interviews from which the ACS estimates are based.² This annual ACS sample is smaller than recent decennial census long-form samples. For example, the Census 2000 long-form sample included about 18 million housing units. The Census 2000 long-form sample was designed to provide reliable estimates for small populations and for small areas such as census tracts and places. The ACS must combine population and housing data from multiple years to produce reliable estimates for geographic areas with populations below 65,000. As a result, the ACS will provide 1-, 3-, and 5-year estimates of data.

ACS data are very timely because they are released in the year immediately following the year in which they are collected. ACS data collected from 2000 through 2004, and published from 2001 through 2005, are available for geographic areas with 250,000 people or more, including all states, the District of Columbia, and many large counties and cities. In 2005, the ACS sample size increased substantially. As a result, starting in 2006, ACS information has been published for areas with populations of 65,000 or more (see Table 2).

² The ACS was not fully implemented until 2005. Prior to 2005 the ACS sampled about 800,000 addresses each year. Information on the initial sample size and the final number of completed interviews in the 2007 ACS are available at <<http://www.census.gov/acs/www/SBasics/SSizes/SSizes07.htm>>.

In December 2008, the first 3-year estimates will be released based on combined data from the 2005, 2006, and 2007 surveys. These estimates will be available for all geographic areas with at least 20,000 people. Three-year estimates will then be updated annually by removing the earliest year and replacing it with the latest year. For example, in 2009, the 3-year estimates will be based on combined data from the 2006, 2007, and 2008 ACS surveys.

By the end of 2009, the ACS will have sampled approximately 15 million addresses, resulting in a total sample large enough to provide data for geographic areas as small as census tracts and block groups. In 2010, the ACS will provide the first 5-year estimates of demographic, housing, social, and economic data for all geographic areas, including those with fewer than 20,000 people.³ These estimates will be based on combined data from the 2005, 2006, 2007, 2008, and 2009 surveys. Like the 3-year estimates, these 5-year estimates will also be updated annually by removing the earliest year and replacing it with the latest one (see Table 2). Thus, in 2011 the 5-year estimates will be based on combined data from the 2006, 2007, 2008, 2009, and 2010 surveys.

³ In both the decennial census and the ACS, all information provided by respondents is confidential. No personal information that could be used to identify individuals or households is ever released or shared. In addition, data will not be published for small geographic areas or groups of people if it would be possible to identify a particular individual.

Table 2. Collection and Release Dates of Data From the ACS
(Pattern repeats after 2012)

| | Release year (late summer or fall) | | | | | | |
|---|---|------|---------------|---------------|---------------|---------------|---------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
| Type of period estimate | Calendar year(s) of monthly data collection | | | | | | |
| 1-year period estimate for areas with 65,000 or more people | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
| 3-year period estimates for areas with 20,000 or more people | | | 2005– 2007 | 2006– 2008 | 2007– 2009 | 2008– 2010 | 2009– 2011 |
| 5-year period estimates for all areas including those with fewer than 20,000 people | | | | | 2005– 2009 | 2006– 2010 | 2007– 2011 |

Source: Citro and Kalton (2007: Table 2-6, page 50).

What Geographic Areas Are Available in the ACS?

The ACS data are tabulated and published for a wide variety of geographic areas. These different types of geographic areas range in size from broad geographic regions (Northeast, Midwest, South, and West) to small towns to census tracts, and even to clusters of city blocks. The population size of an area determines if the ACS data are provided as 1-year, 3-year, or 5-year estimates. Table 3 lists the major types of geographic areas published in the ACS, as well as the type of ACS estimates each area will receive. (Note that Table 3

does not include estimates for Puerto Rico.) For example, ACS data collected in 2006 are currently available for geographic areas with at least 65,000 people, including regions, divisions, states, the District of Columbia, Puerto Rico, congressional districts, Public Use Microdata Areas, and many large counties, metropolitan areas, cities, school districts, and American Indian areas. For a definition and description of these different types of geographic areas, see the glossary at the back of the handbook.

Table 3. Major Geographic Areas and Type of ACS Estimates Available

| 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.

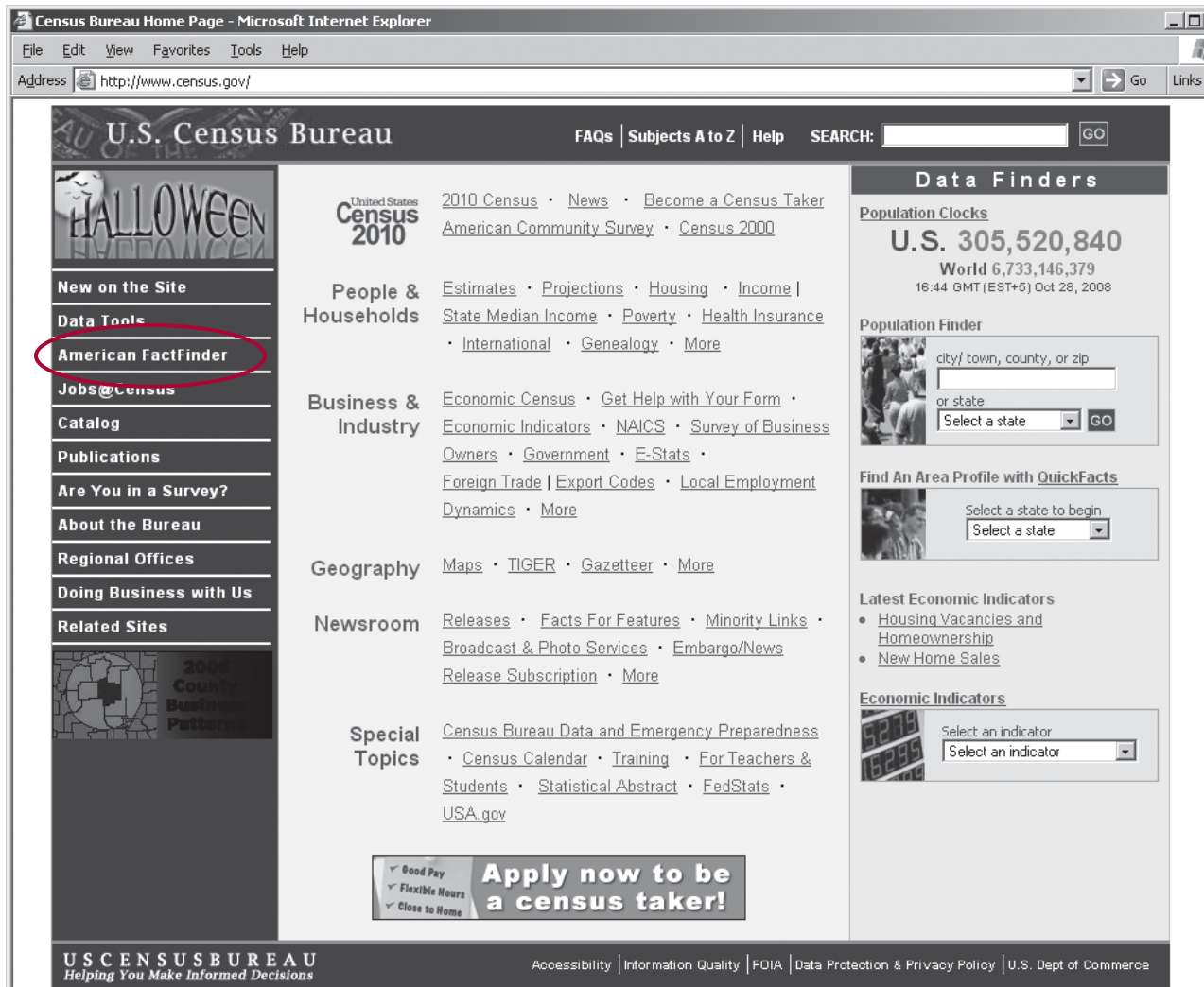
Source: U.S. Census Bureau, 2008. This tabulation is restricted to geographic areas in the United States. It was based on the population sizes of geographic areas from the July 1, 2007, Census Bureau Population Estimates and geographic boundaries as of January 1, 2007. Because of the potential for changes 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.

Accessing ACS Data

ACS data are available through a series of user-friendly products, all of which are available through the Census Bureau's American FactFinder Web site. To get to the American FactFinder Web site, first go to the Census

Bureau's Home Page at <<http://www.census.gov>> and then click on *American FactFinder*, circled in red in the screen shot in Figure 1.

Figure 1. Census Bureau Home Page

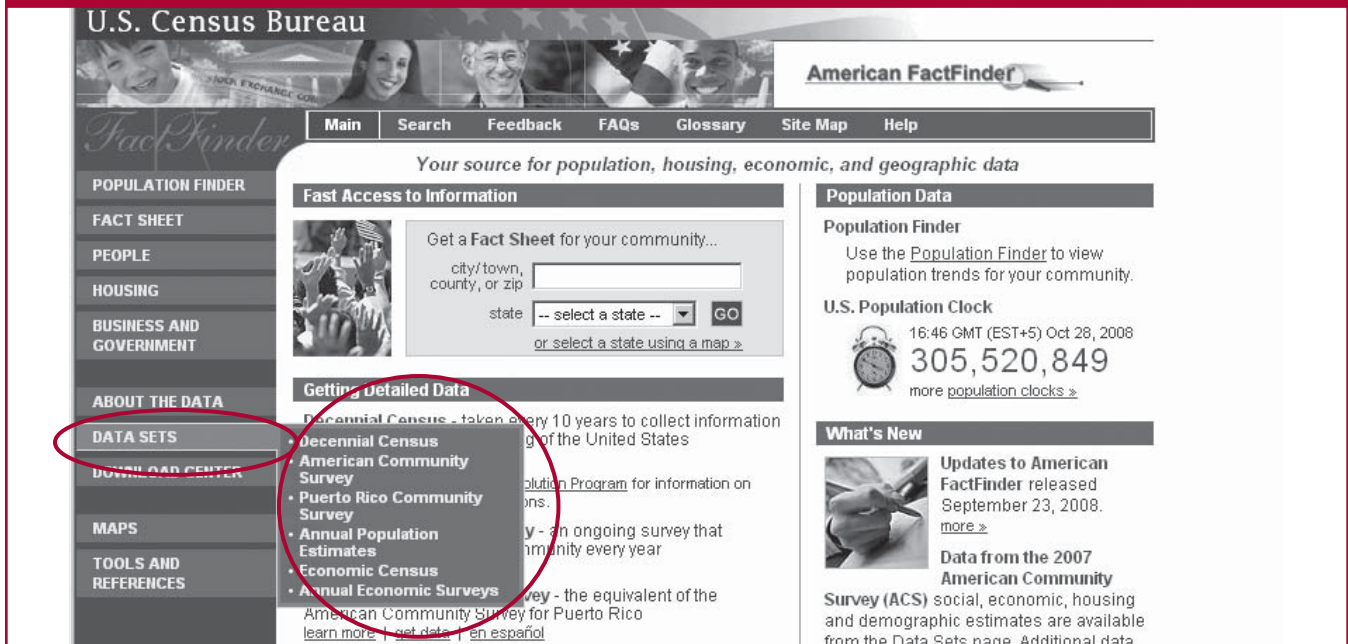


Source: U.S. Census Bureau, American FactFinder, accessed at <<http://factfinder.census.gov>>.

From the American FactFinder home page, click on *Data Sets*, and then choose the *American Community Survey* from the list of options in the drop down menu, as shown in Figure 2. (Note: A tutorial for the

American FactFinder Web site is available online at <http://factfinder.census.gov/home/en/epss/sitetour/3_01.html>.)

Figure 2. American FactFinder Home Page

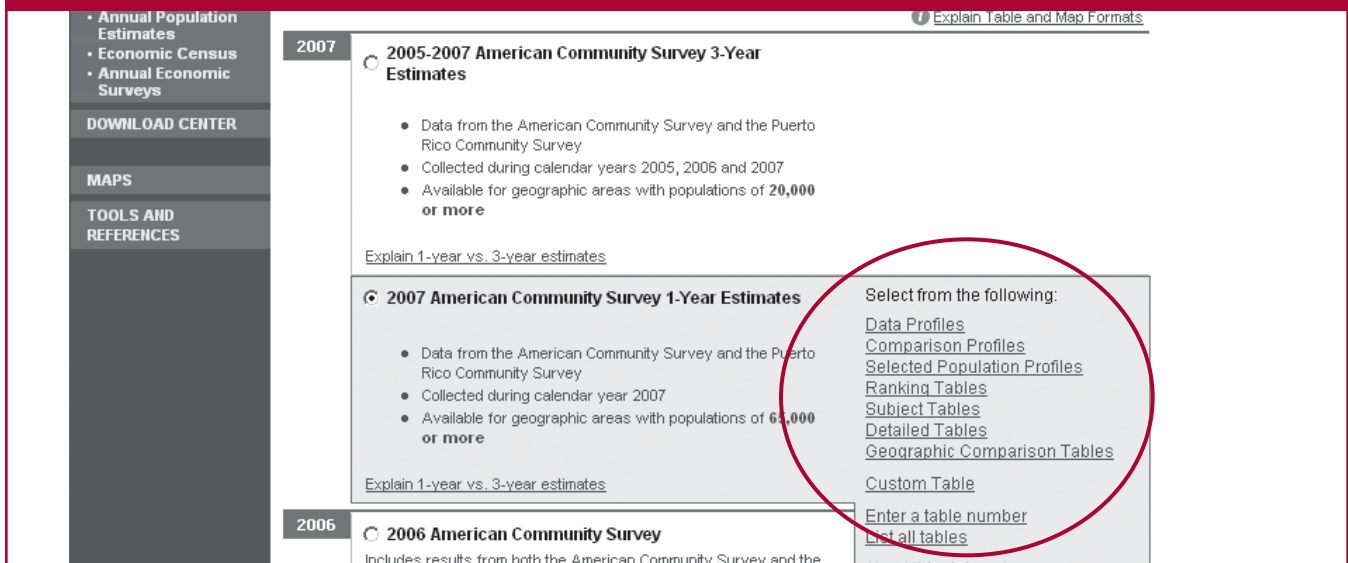


Source: U.S. Census Bureau, American FactFinder, accessed at <<http://factfinder.census.gov>>.

This takes you to a Web page listing all of the data sets available in the American FactFinder for the American Community Survey, from the most recent year back to 2005. Data for the 2000–2004 ACS are archived and only available from the FTP site. The survey for the

most recent year (2007 ACS) is automatically selected, along with the available data products, as shown in Figure 3. ACS data products range from one-page summary tables to detailed cross-tabulations of data.

Figure 3. ACS Data Products Available Through American FactFinder



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

A brief description of each type of ACS data product and the geographic areas it covers is provided in Table 4 below. More detailed information about each data

product is also available on the Census Bureau Web site by clicking on the link *2007 Quick Guide*.

Table 4. **Summary of ACS Data Products**

| Data product | Geographic areas covered | Description |
|-----------------------------------|---------------------------------|---|
| Data profiles | All* | Provide broad social, economic, housing, and demographic profiles. |
| Narrative profiles | All* | Summarize the information in the data profiles using concise, nontechnical text. |
| Selected population profiles | All* | Provide broad social, economic, and housing profiles for a large number of race, ethnic, and ancestry groups. |
| Ranking tables | States, DC, Puerto Rico | Provide state rankings of estimates across 86 key variables. |
| Subject tables | All* | Similar to data profiles but include more detailed ACS data, classified by subject. |
| Detailed tables | All* | Provide access to the most detailed ACS data and cross-tabulations of ACS variables. |
| Geographic comparison tables | All* | Compare geographic areas other than states (e.g., counties or congressional districts) for key variables. |
| Thematic maps | All* | Interactive, online maps that can be used to display ACS data. |
| Custom tables | All* | Allow the user to extract specific rows of data from the ACS detailed tables. |
| Summary files | All* | Provide access to the detailed tables through a series of comma-delimited text files on the Census Bureau's FTP site <www2.census.gov>. |
| Public Use Microdata Sample files | States, DC, Puerto Rico, PUMAs | Provide access to ACS microdata for data users with SAS and SPSS software experience. |

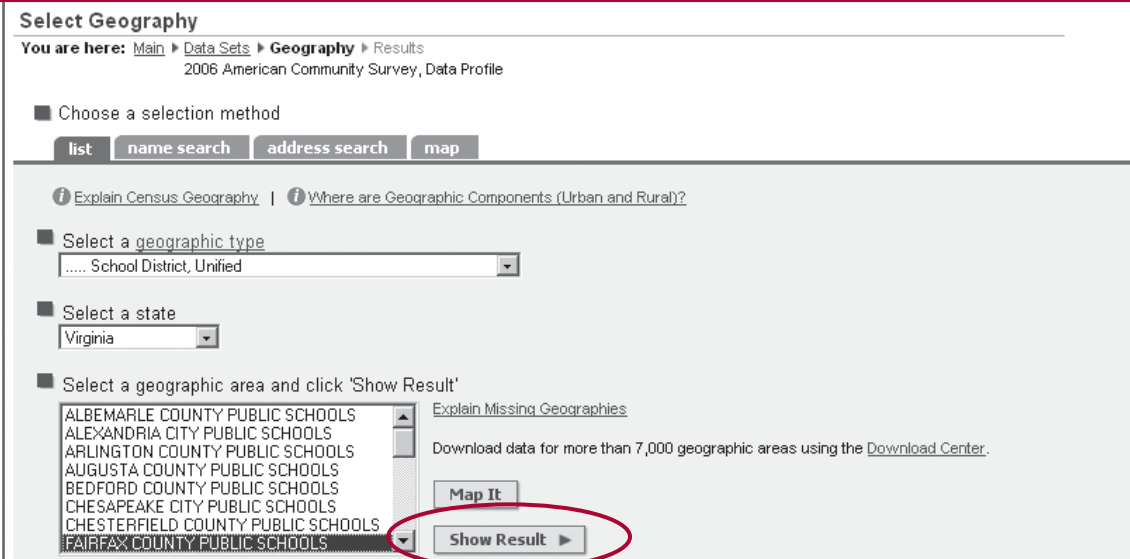
*Note: Data will be available for local areas and small population subgroups with the release of 3-year and 5-year estimates.

Of the different data products, the *Data Profiles* are perhaps the best place to start for high school students because they provide broad social, economic, housing, and demographic profiles of states and local areas. *Data Profiles* provide information ranging from mean travel time to work, to educational attainment, to homeownership rates. After you click on *Data Profiles* (shown in Figure 3), the screen shown in Figure 4 will appear. From the drop-down menus, select the *geographic type*, the *state*, and the specific geographic

area you want. In this example, the geography we have selected is the Unified School District for Fairfax County, Virginia. After making geographic selections, click on *Show Result*, circled in red in Figure 4.

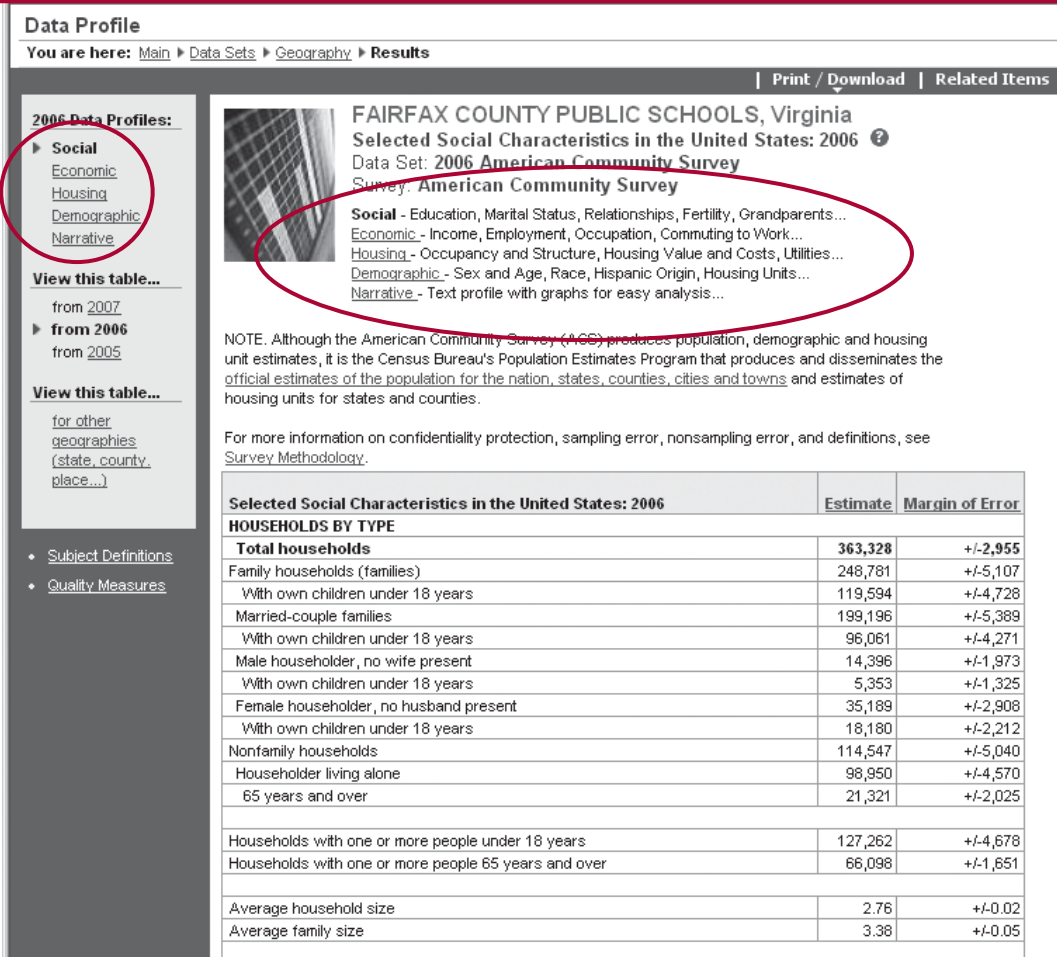
The data shown in Figure 5 will then appear. The *Social Data Profile* pops up as the default profile. To access *Economic, Housing, or Demographic Profiles*, simply click on the links provided, circled in red in Figure 5.

Figure 4. Selecting Geographics in American FactFinder



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

Figure 5. Sample Data Profile for School District

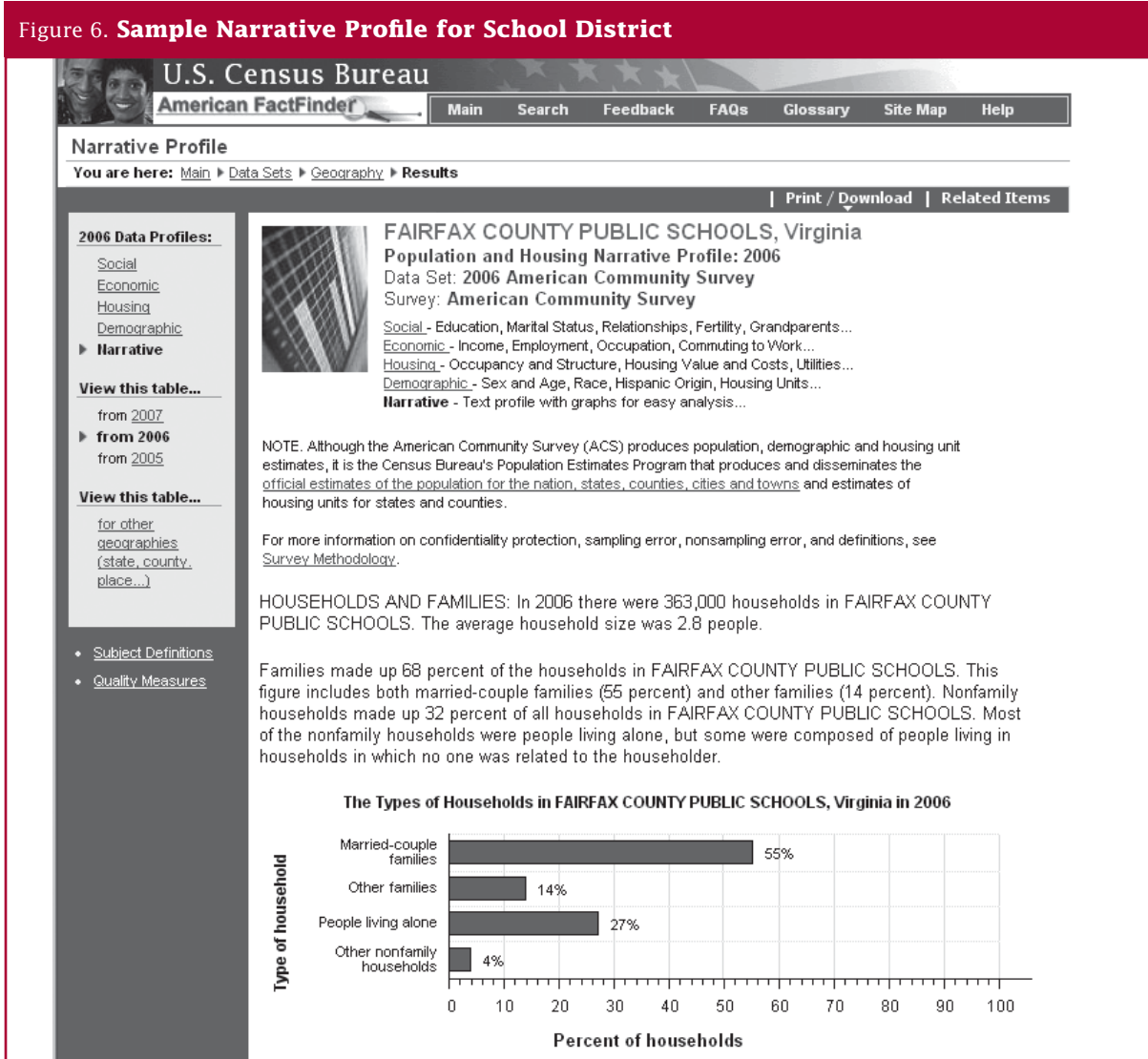


Source: U.S. Census Bureau, American FactFinder, accessed at <http://factfinder.census.gov>.

For those who are just learning how to interpret statistics, *Narrative Profile* summarizes the data in concise, nontechnical terms as illustrated in Figure 6. To access the *Narrative Profile*, just click on the link provided

below the link for the *Demographic Profile*, or use the navigation tool in the left sidebar. Both options are circled in red in Figure 5.

Figure 6. **Sample Narrative Profile for School District**



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

The other ACS data products described earlier can also be accessed through the links provided on the ACS *Data Sets* Web page, as shown in Figure 3.

Ranking Tables provide comparisons of estimates across a set of key variables from the ACS (e.g., median family income or the percentage of people born in Asia). The data can be displayed in either tabular or chart format and are available for the 50 states, the District of Columbia, and Puerto Rico.

The **Geographic Comparison Tables** are similar to the *Ranking Tables* but compare geographic areas *other than states* (e.g., counties or congressional districts) for key variables. The **Thematic Maps** include the same set of variables, but the data are presented in a series

of maps that students can produce through an interactive menu.

Subject Tables are similar to the *Data Profiles* but include more detailed information, classified by subject. For example, students interested in people born outside of the United States can select the table "Selected Characteristics of the Native and Foreign-Born Populations" or another summary table on that topic (see Figure 7). The **Selected Population Profiles** allow the data user to access summary tables for a large number of race, ethnic, and ancestry groups and beginning with the 2007 ACS, summary tables based on country of birth.

Figure 7. Sample Subject Table

U.S. Census Bureau
American FactFinder

Main Search Feedback FAQs Glossary Site Map Help

Subject Tables
You are here: Main > Data Sets > Geography > Tables > Results

Options | Print / Download | Related Items

View this table...
from 2007
from 2006

Change...
geography (state, county, place...)
table

- Subject Definitions
- Quality Measures

United States
S0501. Selected Characteristics of the Native and Foreign-Born Populations
Data Set: 2006 American Community Survey
Survey: American Community Survey

NOTE: For information on confidentiality protection, sampling error, nonsampling error, and definitions, see Survey Methodology.

| Subject | Total | Margin of Error | Native | Margin of Error | Foreign born | Margin of Error | Foreign born; Naturalized citizen | Margin of Error | Foreign born; Not a U.S. citizen | Margin of Error |
|--|--------------------|-----------------|--------------------|-----------------|-------------------|-----------------|-----------------------------------|-----------------|----------------------------------|-----------------|
| Total population | 299,398,485 | **** | 261,850,696 | +/-125,641 | 37,547,789 | +/-125,642 | 15,767,731 | +/-79,241 | 21,780,058 | +/-115,003 |
| SEX AND AGE | | | | | | | | | | |
| Male | 49.2% | +/-0.1 | 49.1% | +/-0.1 | 50.3% | +/-0.1 | 46.8% | +/-0.2 | 52.9% | +/-0.1 |
| Female | 50.8% | +/-0.1 | 50.9% | +/-0.1 | 49.7% | +/-0.1 | 53.2% | +/-0.2 | 47.1% | +/-0.1 |
| Under 5 years | 6.8% | +/-0.1 | 7.7% | +/-0.1 | 0.8% | +/-0.1 | 0.4% | +/-0.1 | 1.2% | +/-0.1 |
| 5 to 17 years | 17.8% | +/-0.1 | 19.3% | +/-0.1 | 7.3% | +/-0.1 | 3.5% | +/-0.1 | 10.0% | +/-0.1 |
| 18 to 24 years | 9.9% | +/-0.1 | 10.0% | +/-0.1 | 9.6% | +/-0.1 | 5.0% | +/-0.1 | 12.9% | +/-0.1 |
| 25 to 44 years | 28.0% | +/-0.1 | 25.7% | +/-0.1 | 43.7% | +/-0.1 | 33.5% | +/-0.2 | 51.0% | +/-0.2 |
| 45 to 54 years | 14.5% | +/-0.1 | 14.2% | +/-0.1 | 16.6% | +/-0.1 | 21.5% | +/-0.2 | 13.0% | +/-0.1 |
| 55 to 64 years | 10.6% | +/-0.1 | 10.6% | +/-0.1 | 10.6% | +/-0.1 | 16.1% | +/-0.1 | 6.5% | +/-0.1 |
| 65 to 74 years | 6.3% | +/-0.1 | 6.3% | +/-0.1 | 6.5% | +/-0.1 | 10.8% | +/-0.1 | 3.4% | +/-0.1 |
| 75 to 84 years | 4.4% | +/-0.1 | 4.5% | +/-0.1 | 3.7% | +/-0.1 | 6.7% | +/-0.1 | 1.6% | +/-0.1 |
| 85 years and over | 1.7% | +/-0.1 | 1.7% | +/-0.1 | 1.3% | +/-0.1 | 2.4% | +/-0.1 | 0.4% | +/-0.1 |
| Median age (years) | 36.4 | +/-0.2 | 35.6 | +/-0.2 | 39.6 | +/-0.2 | 48.3 | +/-0.1 | 34.3 | +/-0.1 |
| RACE AND HISPANIC OR LATINO ORIGIN | | | | | | | | | | |
| One race | 98.0% | +/-0.1 | 97.8% | +/-0.1 | 98.8% | +/-0.1 | 98.5% | +/-0.1 | 98.9% | +/-0.1 |
| White | 73.9% | +/-0.1 | 78.0% | +/-0.1 | 45.3% | +/-0.2 | 44.5% | +/-0.3 | 45.8% | +/-0.3 |
| Black or African American | 12.4% | +/-0.1 | 13.0% | +/-0.1 | 7.8% | +/-0.1 | 8.6% | +/-0.1 | 7.2% | +/-0.1 |
| American Indian and Alaska Native | 0.8% | +/-0.1 | 0.9% | +/-0.1 | 0.4% | +/-0.1 | 0.2% | +/-0.1 | 0.4% | +/-0.1 |
| Asian | 4.4% | +/-0.1 | 1.6% | +/-0.1 | 23.4% | +/-0.1 | 31.0% | +/-0.2 | 18.0% | +/-0.2 |
| Native Hawaiian and Other Pacific Islander | 0.1% | +/-0.1 | 0.1% | +/-0.1 | 0.2% | +/-0.1 | 0.2% | +/-0.1 | 0.3% | +/-0.1 |
| Some other race | 6.3% | +/-0.1 | 4.2% | +/-0.1 | 21.6% | +/-0.2 | 13.9% | +/-0.2 | 27.3% | +/-0.3 |
| Two or more races | 2.0% | +/-0.1 | 2.2% | +/-0.1 | 1.2% | +/-0.1 | 1.5% | +/-0.1 | 1.1% | +/-0.1 |
| Hispanic or Latino origin (of any race) | 14.8% | +/-0.1 | 10.1% | +/-0.1 | 47.2% | +/-0.1 | 30.9% | +/-0.2 | 59.0% | +/-0.2 |
| White alone, not Hispanic or Latino | 66.2% | +/-0.1 | 72.7% | +/-0.1 | 20.5% | +/-0.1 | 28.2% | +/-0.2 | 14.9% | +/-0.2 |
| HOUSEHOLD TYPE | | | | | | | | | | |
| In married-couple family households | 60.4% | +/-0.1 | 60.0% | +/-0.1 | 63.0% | +/-0.2 | 66.7% | +/-0.3 | 60.3% | +/-0.3 |
| In other households | 36.9% | +/-0.1 | 37.1% | +/-0.1 | 35.4% | +/-0.2 | 31.9% | +/-0.3 | 37.9% | +/-0.3 |

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

As the name implies, the **Detailed Tables** include the most detailed ACS data and provide the foundation for other types of data products. *Detailed Tables* are available through the American FactFinder or through the ACS **Summary File** on the Census Bureau's FTP site <www.census.gov/acs/www/Special/acsftp.html>. These *Detailed Tables* may be useful for Advanced Placement students or those working on longer-term projects, but they are generally more useful to researchers. The same is true of the **Public Use Microdata Sample** (PUMS) files, which contain a sample of individual records of people and households that responded to the survey. Note that these records do not contain any personally identifiable information.

Students with access to statistical software packages such as SAS, SPSS, and STATA can use the PUMS files to create custom tabulations of ACS data.

Students may find that some information is not available for some of the more detailed data products, especially for small geographic areas. In some cases, statistics will be missing from a table because there are too few cases to produce a reliable estimate. In other cases, the Census Bureau has suppressed entire tables, because there are too few people in the statistical universe to allow the release or presentation of any data (e.g., characteristics of Navajo Indians living in Maine).

Understanding and Interpreting ACS Data

What Is a Period Estimate?

Because ACS data are collected continuously throughout the year, the ACS provides “period” estimates rather than the “point-in-time” estimates provided by the decennial census. This means that ACS estimates describe the characteristics of the population over a period of time: 12 months for the 1-year estimates; 36 months for the 3-year estimates; and 60 months for the 5-year estimates, whereas Census 2000 estimates approximate a description of the population as of a point in time—April 1, 2000. For most areas with consistent population characteristics over the calendar year, ACS estimates will likely be very similar to those that would be obtained from a point-in-time survey like the decennial census. However, in areas with seasonal populations, such as resort communities, or areas that experience rapid population change during a calendar year, the ACS estimates might be noticeably different from those that would be obtained at one time of the year. ACS estimates also differ from 2000 long-form estimates in several other ways that are described in detail in Appendix 2 of this handbook.

Interpreting and Comparing Single-Year and Multiyear Estimates

While single-year ACS estimates are similar to point-in-time estimates, multiyear estimates are more complex and harder to interpret. For example, it is tempting to describe a 3-year estimate based on data from 2005–2007 as a “2006 average,” since 2006 is the middle year of the period. However, ACS 3-year estimates are not produced by averaging each of the three single-year estimates included in the time period. That is, a 2005–2007 estimate of the percent of children in poverty is not produced by averaging the percents from 2005, 2006, and 2007. Instead, the 2005–2007 estimate is produced by first combining the data records for all 3 years, and then calculating the percent of children in poverty for the 36-month period. The resulting estimate should be interpreted as the “percent of children in poverty in 2005–2007.” Similarly, a 5-year period estimate for 2005–2009 should not be interpreted as a 2007 average. For more information on understanding single-year and multiyear estimates, see Appendix 1 of this handbook.

ACS data users need to be careful *not* to compare different geographic areas using estimates from differing period lengths. A 1-year estimate is not comparable to a 3-year estimate or a 5-year estimate because the time periods are inconsistent. To compare geographic areas, it is important to use a consistent time frame—all 1-year estimates, all 3-year estimates, or all 5-year

estimates. For example, suppose that students wanted to compare estimates for Boston with estimates for Nantucket, a small island off the coast of Massachusetts. Even though the ACS publishes 1-year estimates for Boston, only 5-year estimates will be published for Nantucket. Therefore, in 2010, when 5-year estimates for smaller geographic areas become available, students should compare 2005–2009 estimates for Nantucket with 2005–2009 estimates for Boston, even though more recent single-year estimates are available for Boston. In 2010 and thereafter, the availability of 5-year estimates for every level of geography down to the census tract will allow users to compare data across all geographic areas. Additional guidance on the use and comparison of ACS single- and multiyear estimates is provided in Appendix 1 of this handbook.

ACS data can also be used to look at trends over time. While comparison of single-year estimates for different years is straightforward, comparison of multiyear estimates can be tricky. The Census Bureau recommends comparing estimates for periods that do not overlap, such as comparing 2005–2007 estimates with 2008–2010 estimates. Comparisons of 3-year estimates from 2005 to 2007 and 2006 to 2008 are unlikely to show much difference because two of the years overlap (both include 2006 and 2007 data). The same is true of 5-year estimates with overlapping periods. Finally, as noted earlier, the ACS sample prior to 2006 did not include the group quarters population. Data users should make comparisons of 2006 ACS data with ACS data from prior years only when the geographic area under consideration does not include a substantial group quarters population. Students can also conduct formal tests for statistical significance as described in Appendix 4.

Accuracy of ACS Data

All estimates from sample surveys, including the ACS and the census long-form samples, are subject to error. There are two broad types of error that can occur: sampling error and nonsampling error. Nonsampling errors can result from mistakes in how the data are reported or coded, problems in the sampling frame or survey questionnaires, or problems related to nonresponse or interviewer bias. The Census Bureau tries to minimize nonsampling errors in the ACS by using trained interviewers and by carefully reviewing the survey’s sampling methods, data processing techniques, and questionnaire design. Appendix 6 provides a more detailed description of different types of errors in the ACS and other measures of ACS quality.

Sampling error occurs when data are based on a sample of a population rather than the full population. Sampling error is easier to measure than nonsampling error and can be used to assess the statistical reliability of survey data. For any given area, the larger the sample and the more months included in the data, the greater the confidence in the estimate. The Census Bureau reported the 90-percent confidence intervals for every ACS estimate produced for the 2005 ACS and earlier. Confidence intervals allow data users to assess the level of sampling error, or statistical reliability, of an estimate. A 90-percent confidence interval defines a range around an estimate expected to contain the *true* value with a level of confidence of 90 percent. Beginning with the release of the 2006 ACS data, margins of error are provided for every ACS estimate. Margins of error (MOEs) can be easily converted into confidence intervals. For example, according to 2006 ACS data, 18.3 percent of children under age 18 in the United States live in poverty with a margin of error of plus or minus 0.2 percentage points. By adding and subtracting the MOE from the estimate, we can calculate the

90-percent confidence interval for this estimate. Therefore, we can be 90 percent confident that the *true* child poverty rate in the United States in 2006 falls somewhere between 18.1 percent and 18.5 percent.

18.3% (estimate) – 0.2 percentage points (MOE) = 18.1% (lower confidence limit)

18.3% (estimate) + 0.2 percentage points (MOE) = 18.5% (upper confidence limit)

Detailed information about sampling error and instructions for calculating confidence intervals and margins of error are included in Appendix 3. A classroom application is illustrated later in this handbook.

Now that we have reviewed the types of information the ACS provides and how students can access ACS data on the Census Bureau’s Web site, the next several sections of this handbook provide specific examples to illustrate how teachers in social studies, geography, mathematics, and statistics courses can use ACS in the classroom.

How Teachers Can Use ACS Data

We recognize that there is considerable state-by-state variation in the specific standards and curriculum teachers are required to follow in core subject areas such as mathematics, science, and social studies. Although the United States does not have nationally mandated standards, several professional organizations have developed standards in social studies,

geography, and mathematics.⁴ We have drawn on these standards to illustrate how the ACS can be used in social studies, geography, mathematics, and statistics courses.

Using the ACS in Social Studies Courses

The information gathered in the ACS makes it especially useful for social studies teachers. Many of the topics examined in the ACS can be used to shed light on fundamental institutions and social processes in American society (such as families, fertility, and economic stratification) and a host of social issues (such as poverty, racial equality, and housing).

What makes the ACS particularly appealing for use in the high school classroom is the ability to provide timely information about the communities where students live. For example, the ACS can be used to study the living arrangements of children, the number of elderly people in a community, or the employment situation at the national, state, or local level. The ACS provides a rich set of information about workers, including employment status (working, not working, not in the

labor force) as well as the full-time or part-time status, occupation, and industry of workers. Students can also look at the educational attainment levels of people in particular occupations and the average incomes in those jobs. This information could be particularly useful for students thinking about college and their occupational futures.

The ability to compare data sets from the ACS is also important. For many students, the most meaningful results will come from comparing figures across different geographic areas, comparing population

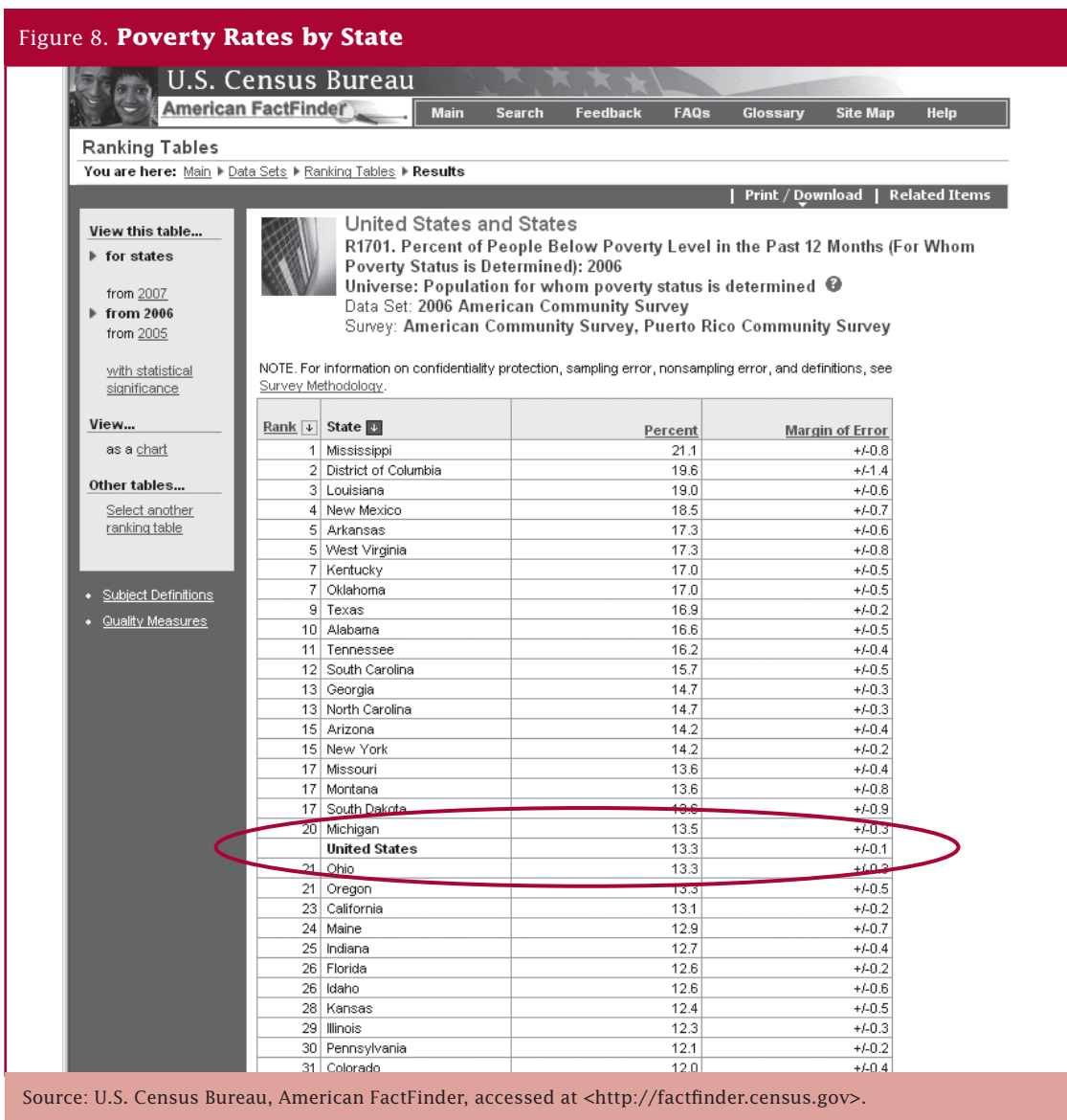
⁴ These include “Geography for Life: The National Geography Standards” from the Geography Education Standards Project; “Expectations of Excellence: Curriculum Standards for Social Studies” from the National Council for the Social Studies; and “Principles and Standards for School Mathematics” from the National Council of Teachers of Mathematics.

subgroups, or looking at trends over time. Sometimes more than one perspective can be included in an analysis, such as comparing change over time between two population subgroups in the same community. For example, one could look at changes in the poverty rates of various population groups over time.

Some examples might help make this perspective clearer. The 2006 ACS shows that the poverty rate (percent of people below the poverty level in the past 12 months) for the state of Ohio is 13.3 percent. For many students, it would be helpful to know if the poverty rate in Ohio is higher or lower than the poverty

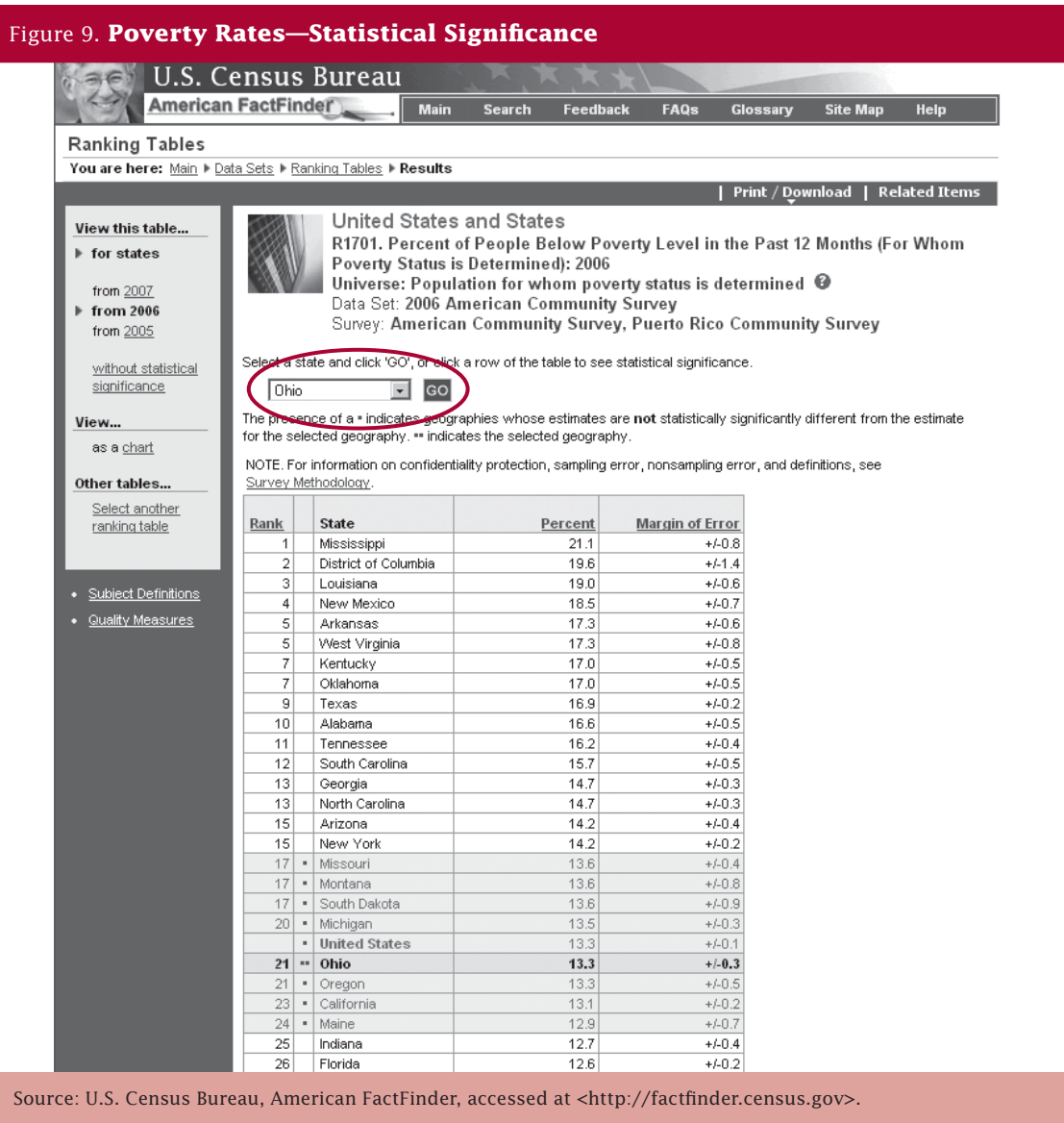
rate for another state, for example, Michigan. Since the ACS provides consistent data for cities and states, it is easy to make this comparison. The 2006 ACS indicates that the poverty rate for Ohio is 13.3 percent and for Michigan it is 13.5 percent. Students may observe that the poverty rate in Ohio appears to be lower than the poverty rate for Michigan. But they cannot draw that conclusion without looking more closely at these data.

To provide additional perspective, students can access an ACS *Ranking Table* to see how poverty in Ohio compares with poverty levels in other states (see Figure 8).



This ACS table shows that Ohio ranks twenty-first among the states in terms of the poverty rate and students can see that Michigan ranks twentieth. (Note that a rank of 1 indicates the highest estimated poverty rate, not the best poverty rate.) Students can also see from this ranking function that the poverty rate among states varies from a low of 7.8 percent in Maryland (not

visible in Figure 8) to a high of 21.1 percent in Mississippi. Clicking on the “with statistical significance” link in the left sidebar provides important results of testing to determine if the estimates for these two states are different beyond sampling error. Figure 9 shows that the poverty rate in Ohio is not statistically different from the rate for Michigan.



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

Students may also be interested in using the ACS **Subject Tables** to explore how different groups in a city such as Columbus, Ohio, compare with each other in terms of poverty (see Figure 10). *Subject Tables* is also accessed from the link shown in Figure 3. After you click on *Subject Tables*, a geography selection screen will pop up as illustrated previously in Figure 4.

This ACS *Subject Table* shows that the poverty rate for the White alone, not Hispanic or Latino population in

Columbus is 14.6 percent compared with 33.1 percent for Blacks or African Americans and 25.9 percent for people of Hispanic or Latino origin. One lesson from this type of “decomposition” is that the overall poverty rate in Columbus, Ohio, masks very different rates for subgroups such as Whites, Blacks, and Latinos. Such decomposition could focus on other characteristics such as age, sex, or home ownership. The breadth of data in the ACS enables teachers to customize their lesson plans based on local issues.

This table also illustrates one of the challenges in using ACS data for smaller geographic units, as described earlier. Note that the poverty rate is not shown for American Indians and Alaska Natives, Asians, Native Hawaiians and Other Pacific Islanders, Some Other Race, or Two or More Races. This is because there were not sufficient numbers of people in these racial groups in Columbus for the Census Bureau to provide reliable estimates. The prevalence of missing data in some tables will be reduced when multiyear estimates become available starting in December 2008.

The ACS can also be used to examine changes over time. As noted above, the 2006 poverty rate in Ohio was 13.3 percent (with a *margin of error* of 0.3 percent). But what was it in earlier years? The ACS shows that in 2005 the poverty rate in Ohio was 13.0 percent (with a *margin of error* of 0.3 percent). This comparison suggests that poverty has increased in Ohio between 2005 and 2006, but students would need to conduct a statistical test to determine if the estimate for 2006 is actually statistically different from the estimate for 2005. For more information about conducting

such statistical tests, see Appendixes 3 and 4 at the end of this handbook.

These are just a few examples of some of the important issues teachers can have students address with data from the ACS. Others could include such topics as:

- What is the average (mean) income of families in your community compared with the average for the state?
- What percentage of the population in your state is not working or looking for work?
- What is the median value of houses in your community?
- How many households in your county have no vehicles?
- What is the high school dropout rate (percent of 16-to-19-year-olds who are not in school and do not have a high school diploma) in your city?

Figure 10. Poverty Rates Among Subgroups in Columbus, Ohio

U.S. Census Bureau
American FactFinder

Subject Tables
You are here: Main > Data Sets > Geography > Tables > Results

Columbus city, Ohio
S1701. Poverty Status in the Past 12 Months
Data Set: 2006 American Community Survey
Survey: American Community Survey

NOTE: For information on confidentiality protection, sampling error, nonsampling error, and definitions, see Survey Methodology.

| Subject | Total | Margin of Error | Below poverty level | Margin of Error | Percent below poverty level | Margin of Error |
|---|----------------|------------------|---------------------|-----------------|-----------------------------|-----------------|
| Population for whom poverty status is determined | 699,167 | +/-10,276 | 144,832 | +/-9,160 | 20.7% | +/-1.3 |
| AGE | | | | | | |
| Under 18 years | 172,103 | +/-5,354 | 49,556 | +/-5,505 | 28.8% | +/-3.0 |
| Related children under 18 years | 170,713 | +/-5,458 | 48,298 | +/-5,483 | 28.3% | +/-3.0 |
| 18 to 64 years | 466,144 | +/-7,308 | 88,675 | +/-5,261 | 19.0% | +/-1.1 |
| 65 years and over | 60,920 | +/-1,977 | 6,601 | +/-1,235 | 10.8% | +/-2.0 |
| SEX | | | | | | |
| Male | 344,398 | +/-6,085 | 65,636 | +/-5,132 | 19.1% | +/-1.4 |
| Female | 354,769 | +/-5,984 | 79,196 | +/-5,702 | 22.3% | +/-1.6 |
| RACE AND HISPANIC OR LATINO ORIGIN | | | | | | |
| One race | N | N | N | N | N | N |
| White | 449,903 | +/-9,893 | 67,416 | +/-5,768 | 15.0% | +/-1.2 |
| Black or African American | 195,372 | +/-4,898 | 64,709 | +/-7,161 | 33.1% | +/-3.5 |
| American Indian and Alaska Native | N | N | N | N | N | N |
| Asian | N | N | N | N | N | N |
| Native Hawaiian and Other Pacific Islander | N | N | N | N | N | N |
| Some other race | N | N | N | N | N | N |
| Two or more races | N | N | N | N | N | N |
| Hispanic or Latino origin (of any race) | 28,612 | +/-2,445 | 7,423 | +/-1,948 | 25.9% | +/-6.5 |
| White alone, not Hispanic or Latino | 432,095 | +/-9,211 | 63,278 | +/-5,225 | 14.6% | +/-1.2 |
| EDUCATIONAL ATTAINMENT | | | | | | |
| Population 25 years and over | 447,121 | +/-6,598 | 64,945 | +/-4,412 | 14.5% | +/-1.0 |
| Less than high school graduate | 60,892 | +/-3,479 | 18,949 | +/-2,772 | 31.1% | +/-3.7 |

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

The ACS can also be used to address several specific social science standards as specified by the National Council for the Social Studies (NCSS).⁵ We briefly

describe several of these standards in Table 5 and suggest how ACS data can be used to address each standard.

Table 5. Examples of Using ACS Data in the Context of Standards From the National Council for the Social Studies

| Standards for the national council for social studies | Questions and topics that can be examined with ACS data | Data elements in the ACS that pertain to this question or issue |
|--|---|--|
| Culture and cultural diversity | Describe cultural similarities and differences between the local community and the state or nation. | Race, Hispanic origin, ancestry, place of birth, language spoken at home |
| People, places, and environments | Describe links of people in your community with other places. Compare the population density of your community with that of other communities in the state. | Place of birth, language spoken at home, residence one year ago, population density |
| Production, distribution, and consumption | Describe the workforce in the local community. | Occupation, industry, class of worker |
| Science, technology, and society | Describe the kinds of scientists who work in your community. | Occupation, industry, class of worker |
| Global connections | Describe the people in your community or state who were born outside the United States. | Place of birth, language spoken at home, ability to speak English |
| Civic ideals and practices | Describe the potential voters in your community. | Number and characteristics of people over age 18, number and characteristics of citizens over age 18 |

Source: U.S. Census Bureau, based on the National Standards for Social Studies Teachers.

Using the ACS in Geography Courses

The American FactFinder Web site also provides students with the capability to interactively create thematic maps using ACS data. Teachers can use ACS data and the American FactFinder to address several of the National Geography Standards, such as understanding the characteristics, distribution, and migration of human populations on Earth's surface, and using maps and other geographic representations, tools, and technologies to acquire, process, and report information from a spatial perspective.⁶

Students can explore the spatial distribution of immigrants across the United States using ACS data on the percent of foreign-born individuals by state. Students can use *Thematic Maps* in the American FactFinder to create a choropleth map by state (see Figure 11).⁷ *Thematic Maps* is accessed through the link provided on the ACS *Data Sets* page as shown in Figure 3. After you click on *Thematic Maps*, the *Select Geography* screen will appear. For maps, you can either select "nation" to get a map of the entire United States or

⁵ National Council for the Social Studies, 2002, National Standards for Social Studies Teachers, revised 2002.

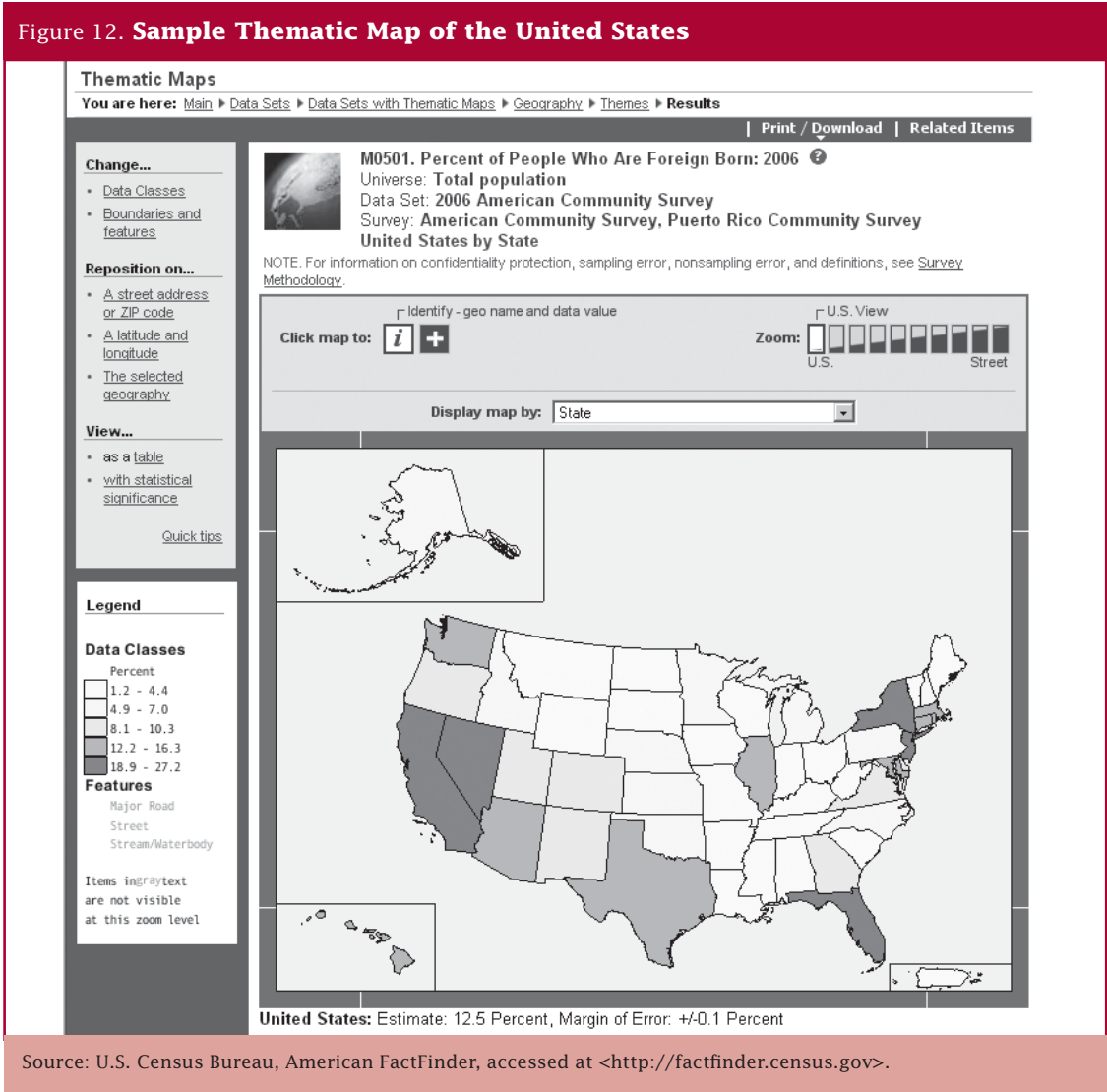
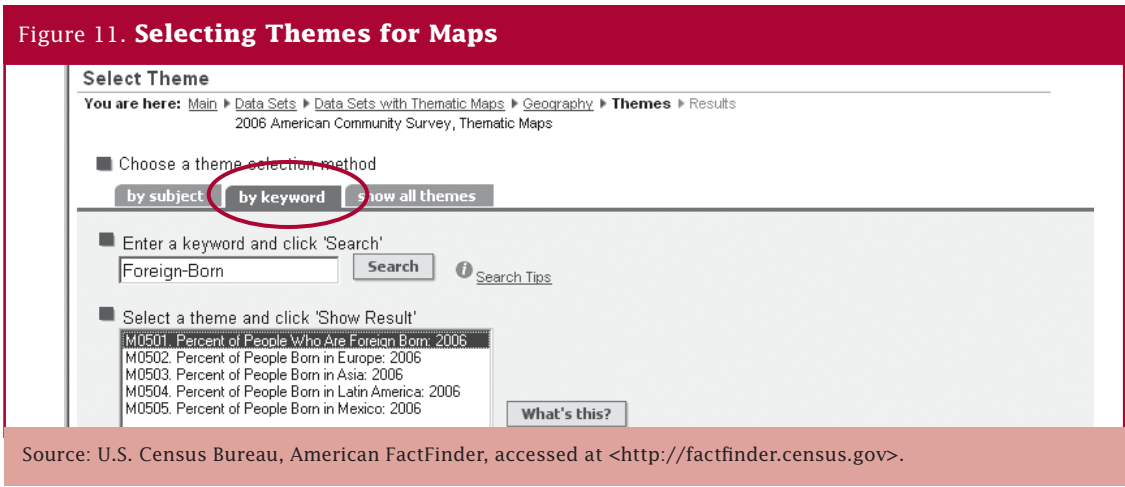
⁶ Geography Education Standards Project. 1994. *Geography for Life: The National Geography Standards*. Washington, DC: National Geographic Society Committee on Research and Exploration.

⁷ A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the characteristic that is being displayed on the map.

“state” to get a map of just one state. Once you select a geography, the *Select Theme* screen will appear as shown in Figure 11. To find the data you want, you can search by subject or keyword or see a list of all possible themes or tables. In this example, we search by the

keyword “foreign-born,” and select the theme *Percent of People Who Are Foreign Born: 2006*.

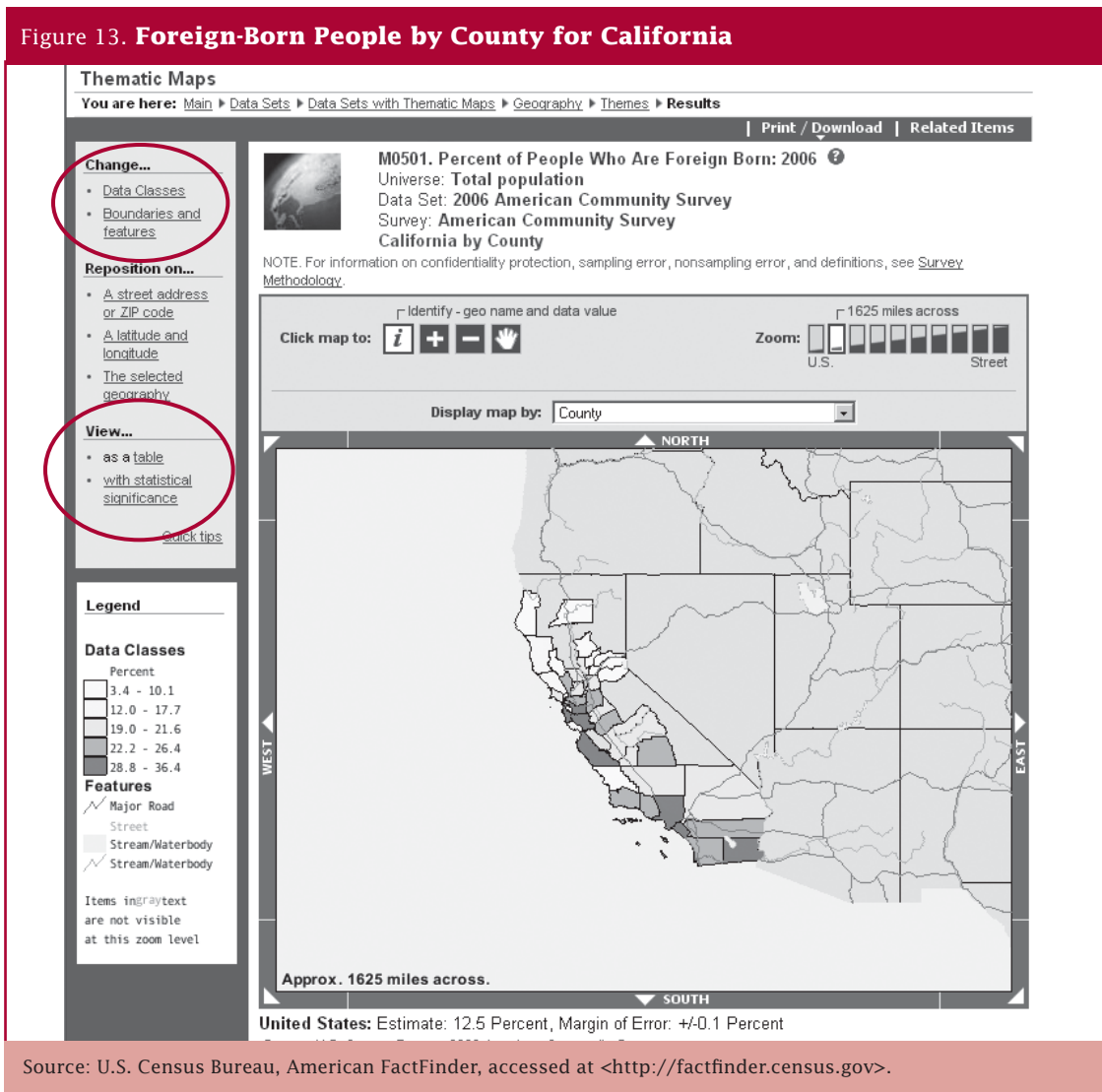
With *nation* selected as the geography, the thematic map is shown in Figure 12.



This map quickly reveals the states with the highest proportion of immigrants including California, Nevada, Florida, New Jersey, and New York, where 18.9 percent or more of the population was born outside of the United States. Students could also use ACS data to drill down within a particular state like California to examine how immigrants are spatially distributed across counties (see Figure 13). They would do this by selecting *state* instead of *nation* in the *Select Geography* screen. Of course, as explained above, single-year estimates are available only for counties with a population of at least 65,000. That is why some California counties are not shown. Once 3-year estimates for 2005–2007 are

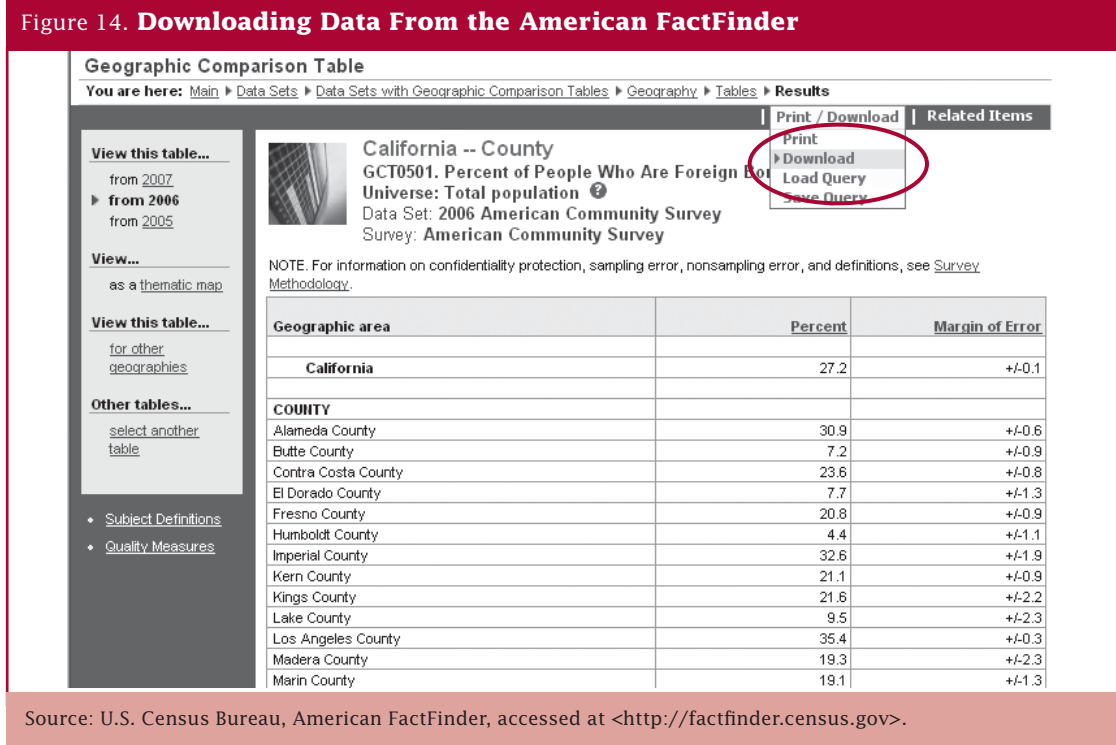
available in December of 2008, data will be shown for all counties with populations of at least 20,000.

Based on the county map, students discover that immigrants are distributed throughout California, but are most concentrated in counties like Los Angeles and Imperial, where the share who are foreign born can reach as high as 36 percent. Although the thematic maps in the American FactFinder pop up with the ranges or data classes and color scheme automatically selected, the navigation tools in the left sidebar (circled in red in Figure 13) provide options for students to change the *Data Classes* and color scheme.



If students click on the *View as a table* navigation tool in the left sidebar (circled in red in Figure 13), a screen will pop up with a table showing the data for every state (see Figure 14). Students can download data tables like this one in Microsoft Excel format to

import for use with other mapping software such as ArcView or MapViewer. To do this, they simply click on the *Download* option as highlighted in Figure 14, and a window will pop up with download options, including Microsoft Excel.

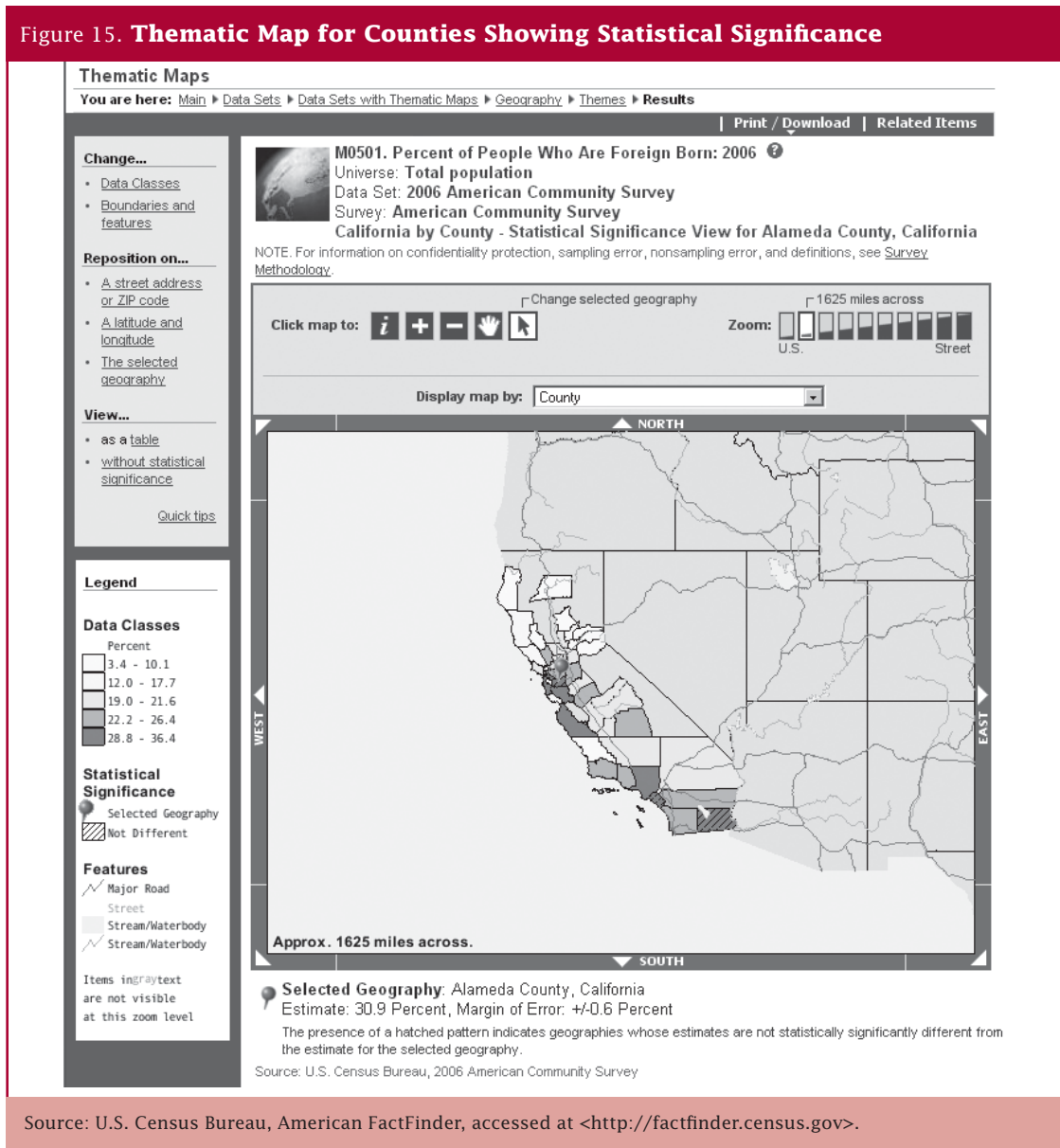


If students click on the navigation tool *View with statistical significance* (also circled in Figure 13), the thematic map will change to indicate with cross-hatching which counties have an estimate of percent foreign born that is **not** statistically different from the percentage in a reference county (see Figure 15).

With Los Angeles County selected as the reference county, this map indicates that the estimates for every other county are statistically different from those for Los Angeles County. This tool provides a quick way for students to assess statistical significance without having to calculate it themselves.

The wealth of data in the ACS combined with this easy-to-use thematic mapping capability offer a number of opportunities for geographic activities on the American FactFinder Web site. Students can also download a wide variety of ACS data tables for use with other mapping software applications. However, students who download such data tables will need to be aware of missing geographic areas when they import these tables for use with other mapping packages. For example, if they download a 2006 ACS table with county-level data for California, the resulting Excel file will be missing data for all counties with fewer than 65,000 people.

Figure 15. Thematic Map for Counties Showing Statistical Significance



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

Using the ACS in Mathematics and Statistics Courses

Mathematics and statistics teachers have a multitude of data sources they can use in developing homework problems, project assignments, and examples for course lectures. The ACS offers three unique advantages as a potential data source:

1. The ACS provides a rich source of current data about students' local communities, their state, and the nation. Students may find it easier to understand and remember mathematical and statistical concepts and skills they learn with data and examples they find both interesting and personally relevant.

2. ACS data are easily accessible online and are updated annually. This saves teachers the work of creating and maintaining data sets each year for students to use for assignments and projects.

3. The ACS estimates provided online include their margins of error, making it easy for teachers and students to use these data to calculate additional measures of error (e.g., standard errors or coefficients of variation) and to conduct tests of statistical significance.

In selecting examples to illustrate potential applications of ACS data in math and statistics courses, we consulted the *Principles and Standards for School Mathematics*, published by the National Council of Teachers of Mathematics (NCTM), as well as the 2007–2008 detailed course description for AP Statistics from the College Board.^{8,9} Our review suggests that ACS data and information can be used to address components of several of the NCTM standards, including Measurement and Data Analysis and Probability, as well as the section on Statistical Inference of the AP statistics course. Teachers and students in AP statistics may also be interested in more detailed information about the sample design and methodology of the ACS. A technical paper is available online at <<http://www.census.gov/acs/www/Downloads/tp67.pdf>>.

Proportions and Ratios

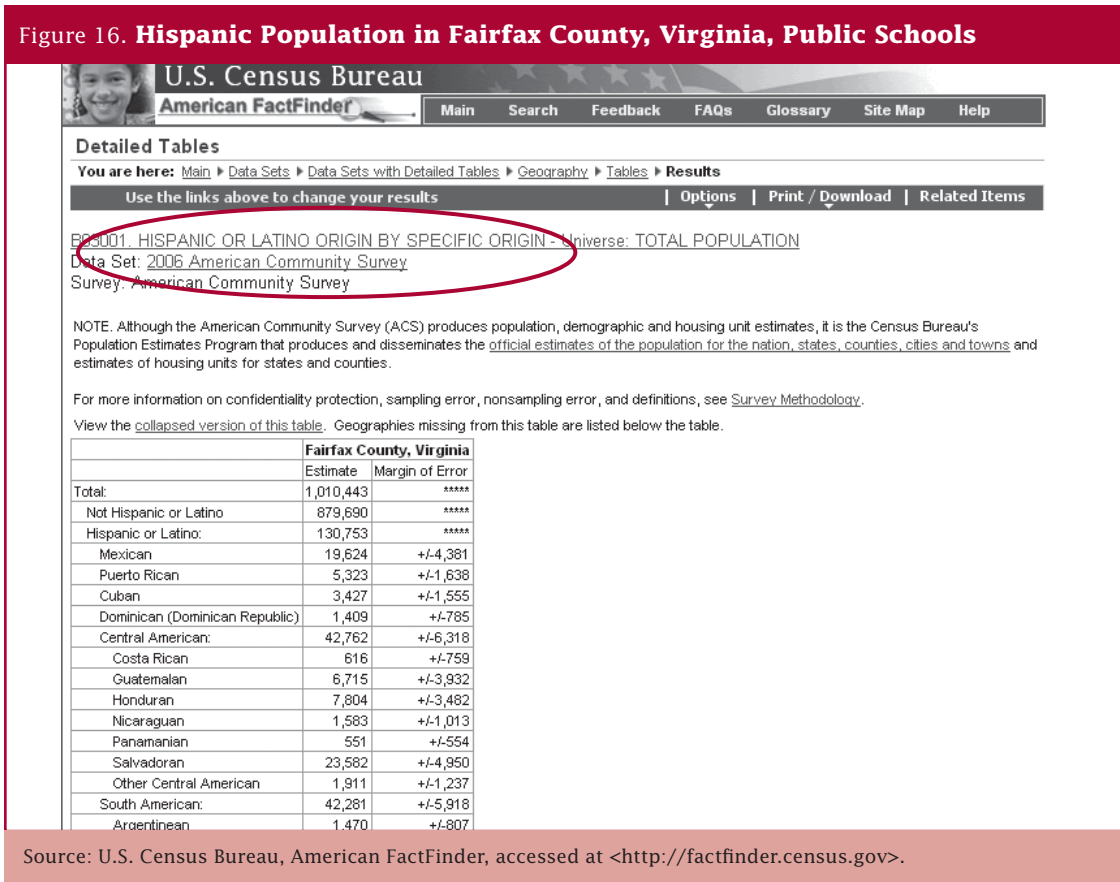
ACS data can be used to teach students how to calculate proportions and ratios for a variety of topics for their school district, city, county, or state. For example, to analyze the Hispanic population of their school district, students in Fairfax County, Virginia, could access

the table “Hispanic or Latino Origin by Specific Origin” from the ACS product *Detailed Tables* (see Figure 16).

The data in this table could be used to calculate the proportion of the population in this school district that is Hispanic (e.g., 130,753/1,010,443) or the proportion of the Hispanic population that is Mexican (e.g., 19,624/130,753). It could also be used to calculate the ratio of Central Americans to Mexicans within the school district (e.g., 42,762/19,624).

Statistical Inference

As described previously in this handbook, all estimates provided in ACS online data products either include the *margin of error* or the upper and lower bounds of a 90-percent *confidence interval*. Another measure of the precision of an estimate is the *standard error*. Standard errors measure the variability of an estimate due to sampling, and they are needed to conduct tests of statistical significance. Statistical significance indicates whether the difference between two estimates is likely to represent a real difference that exists within the full population or whether instead it has occurred by chance due to sampling. In this example, we illustrate



⁸ National Council of Teachers of Mathematics, *Principles and Standards for School Mathematics*, Reston, Virginia: 2000.

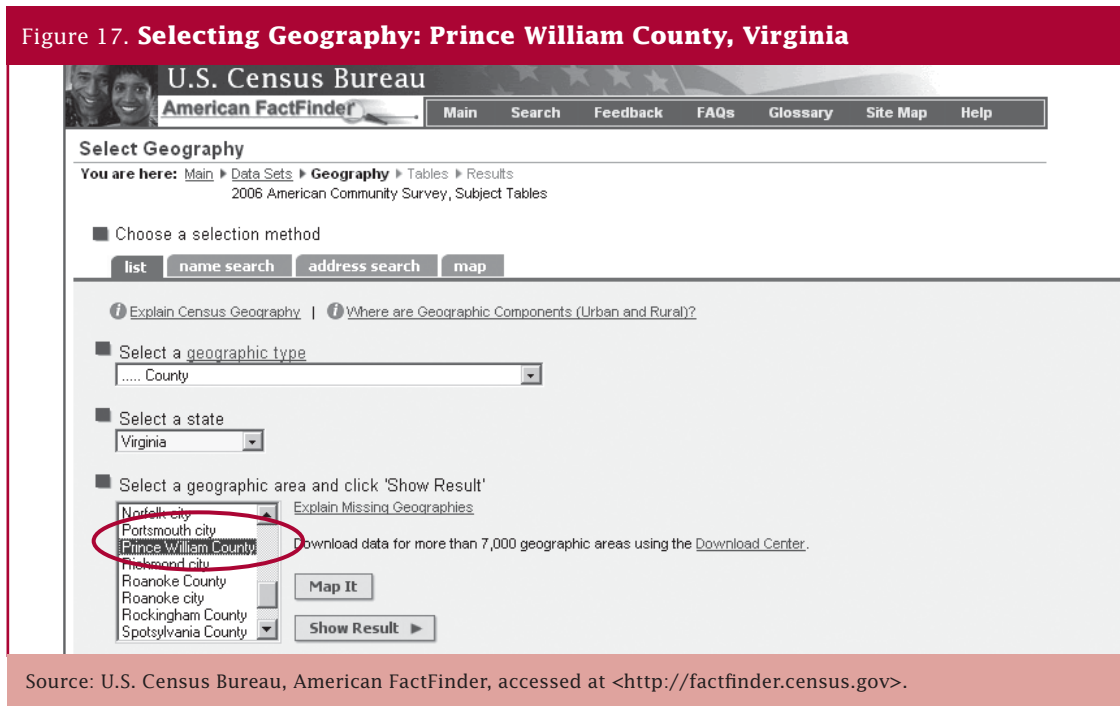
⁹ College Board, *AP Statistics Course Description: May 2007–2008*.

how teachers and students can use the MOEs for ACS estimates to calculate additional MOEs and to test for statistical significance on a topic related to current events. For more information on how to calculate and interpret measures of sampling error, see Appendix 3 at the end of this handbook.

In the last year, foreclosures and falling housing values have been widely reported. However, prior to 2007, real estate markets in many metropolitan areas were still booming, and students can use the ACS to track the changes in home values in their communities. In this example, we'll assume that students in Prince William County, Virginia, want to determine if their

county experienced a statistically significant increase in housing values between 2005 and 2006.¹⁰ Because they want to establish a baseline for homes that could potentially be affected by rising interest rates and foreclosure, they only want to examine homes that have a mortgage, not those that are owned free and clear.

To find ACS data on housing values, students can check the *Subject Tables* found on the *Data Sets* page on the American FactFinder (AFF) Web site (shown in Figure 3). When they click on *Subject Tables*, the *Select Geography* screen shown in Figure 17 appears. After making the selections shown below to get data for Prince William County, they click *Next*.



¹⁰ The 2005 ACS survey did not include group quarters population, while the 2006 ACS did. Therefore, as noted earlier in this handbook, students need to be careful when comparing data estimates from 2005 and 2006. In this example, the data item of interest is collected only for the household population, not the group quarters population. Therefore, it is reasonable to make comparisons between 2005 and 2006 ACS estimates for this data item.

On the main *Subject Tables* screen students can use the drop-down menu to search tables by major topic areas (see Figure 18). For this example, they would select the subject *Housing Financial Characteristics* to get the three table choices shown in Figure 19.

By selecting Table S2506, *Financial Characteristics for Housing Units With a Mortgage*, they obtain the distribution of owner-occupied housing units by value for 2006.

Figure 18. Searching Subject Tables by Major Topic

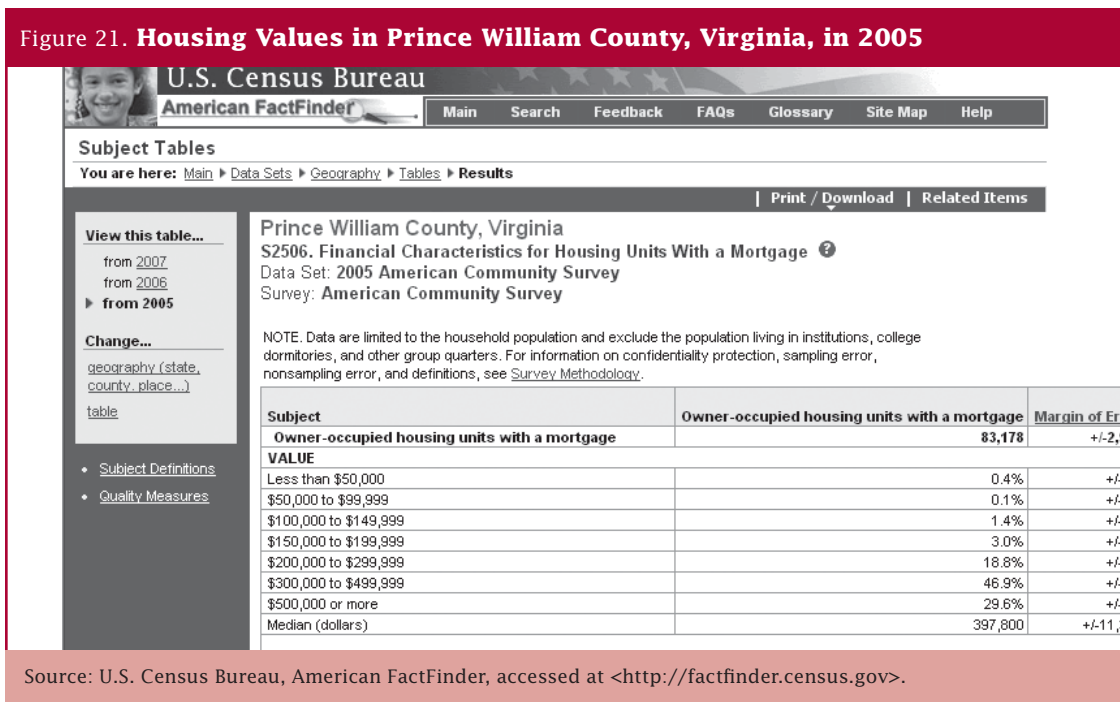
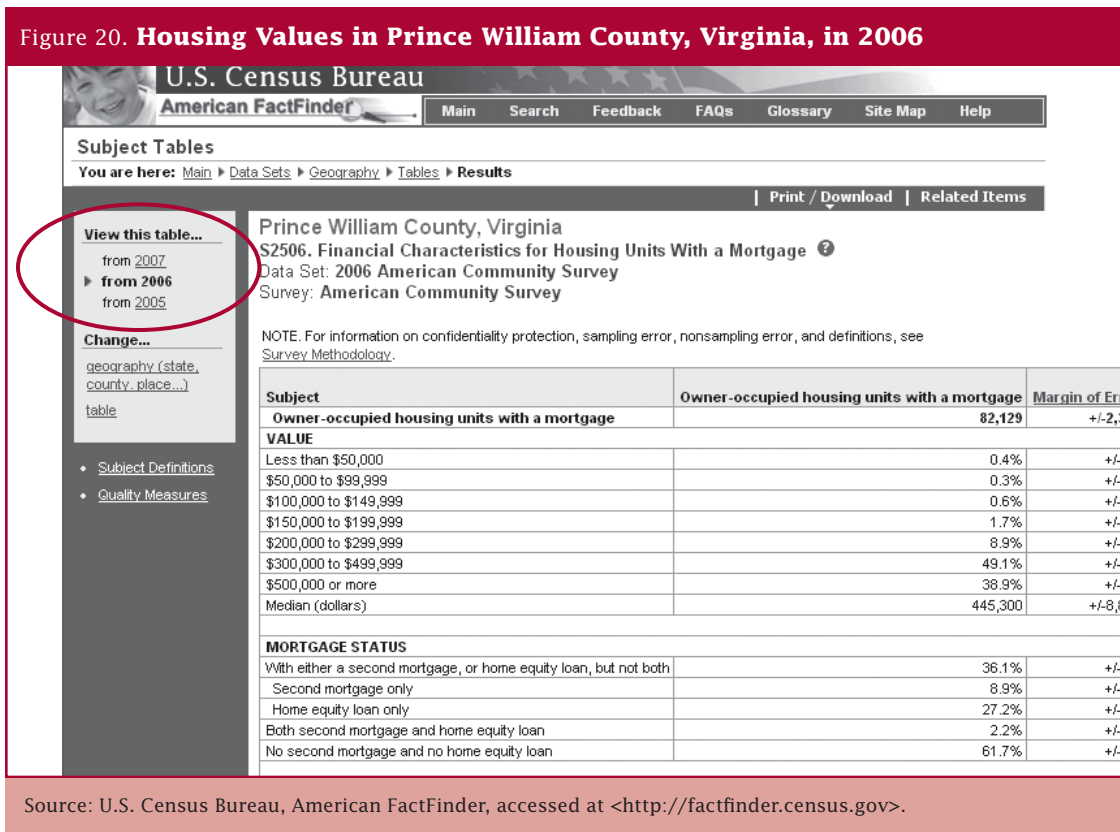
The screenshot shows the 'Subject Tables' page on the American FactFinder website. The page title is 'Subject Tables' and the breadcrumb trail is 'Main > Data Sets > Geography > Tables > Results'. A dropdown menu is open, showing a list of subject categories: Disability, Education, Employment, Fertility, Grandparents, Group Quarters Population, Housing, Housing Financial Characteristics (highlighted with a red circle), Housing Physical Characteristics, Income, and Mobility. The 'GO' button is visible next to the dropdown. Below the dropdown, there are several subject table entries under different categories: 'Age and Sex' (S0101 Age and Sex), 'Aging' (S0102-PR Population, S0102 Population 60 and Over, S0103-PR Population, S0103 Population 65 Years and Over in the United States), and 'Children' (S0901 Children Characteristics, S1001 Grandchildren Characteristics, S0902 Teenagers' Characteristics). The source is cited as: U.S. Census Bureau, American FactFinder, accessed at <http://factfinder.census.gov>.

Figure 19. Finding Tables With Data on Housing Values

The screenshot shows the 'Subject Tables' page on the American FactFinder website. The dropdown menu is now set to 'Housing Financial Characteristics'. The 'GO' button is visible next to the dropdown. Below the dropdown, there are three subject table entries under the 'Housing Financial Characteristics' category: S2503 Financial Characteristics, S2506 Financial Characteristics for Housing Units With a Mortgage, and S2507 Financial Characteristics for Housing Units Without a Mortgage. The source is cited as: U.S. Census Bureau, American FactFinder, accessed at <http://factfinder.census.gov>.

By clicking on the *View this table from 2005* navigation tool in the left sidebar (circled in red in Figure 20), stu-

dents can quickly and easily get the comparable data for Prince William County for 2005 (see Figure 21).



These data indicate that the median value of owner-occupied homes with a mortgage increased from \$397,800 in 2005 to \$445,300 in 2006 in this county. These data also show an increase from 29.6 percent to 38.9 percent in the share of homes with values of \$500,000 or more. It appears that this shift of homes into the highest value category may be resulting in the jump in median home value from 2005 to 2006. Using the *margin of error* provided in the tables, students can conduct a statistical test to determine whether the change from 2005 to 2006 in the share of homes in the top value category is statistically significant.

To determine whether the observed difference between 29.6 percent in 2005 and 38.9 percent in 2006 is statistically significant, students must conduct a test using the following formula:¹¹

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

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 ($i=1,2$)

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

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).¹²

First, students need to derive the standard error for each of these estimates using the following formula:

$$SE = \frac{MOE}{1.645}$$

Substituting the values for 2006 yields:

$$SE = \frac{2.3}{1.645} = 1.398$$

The standard error for the 2005 estimate is:

$$SE = \frac{2.6}{1.65} = 1.576^{13}$$

¹¹ Step-by-step instructions for this calculation as well as general guidance on making comparisons with ACS data are included in Appendix 4 at the end of this handbook. Instructions for calculating the MOE, confidence interval, standard error, and coefficient of variation, are detailed in Appendix 3.

¹² The MOEs published with the ACS estimates correspond to the 90-percent confidence level. If students want to conduct a test at the 95-percent or 99-percent confidence level, then the published MOEs must be converted to the desired confidence level by using an adjustment factor. Detailed instructions for making these adjustments are provided in Appendix 3 of this handbook.

¹³ Note that for ACS data from 2005 and earlier years, the value 1.65 was used to derive the MOE, and therefore is used in the denominator in this calculation rather than 1.645, which was used beginning with the 2006 ACS data.

Filling in the values for the statistical test for this example results in the following:

$$\left| \frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}} \right| = \left| \frac{38.9 - 29.6}{\sqrt{(1.398)^2 + (1.576)^2}} \right| = \left| \frac{9.3}{\sqrt{1.954 + 2.484}} \right| = \left| \frac{9.3}{\sqrt{4.438}} \right| = \left| \frac{9.3}{2.107} \right| = 4.414$$

Since the test value (4.41) is greater than the critical value for a confidence level of 90 percent (1.645), then students could conclude that the percent of owner-occupied homes with values of \$500,000 or more in their county had indeed increased significantly between 2005 and 2006.

Data users can also examine the *relative* amount of sampling error associated with a particular estimate by calculating a coefficient of variation (CV). CVs relate the size of the standard error to the size of the estimate itself. Larger CVs indicate less reliable estimates. A general guideline is that CVs should be less than 30 percent for estimates to be considered reliable. For the example above, the CV for the 2005 estimate of the percent of homes with values of \$500,000 or more is:

$$CV = \frac{1.576}{29.6} = 0.0532 * 100 = 5.3\%$$

The CV for the 2006 estimate is:

$$CV = \frac{1.398}{38.9} = 0.0351 * 100 = 3.5\%$$

Since both of these CVs are below 10 percent, these estimates can be considered to be reliable.

The availability of ACS data for multiple years, time periods (2005–2007, 2005–2009), and geographic areas offers students extensive choices in making comparisons. However, it is important to follow the general guidelines for interpreting single- and multi-year estimates when making comparisons and conducting tests of statistical significance. In addition, due to changes in geographic boundaries, sample frame, and collection methods in the ACS over time, students may find the guidance in Appendix 4 of this handbook helpful in determining whether a particular comparison is appropriate and valid.

Other Resources for Working With ACS Data

Teachers and students can find a wealth of additional information about the ACS on the Census Bureau's Web site <www.census.gov/acs>. However, sorting through all of the information available on the ACS would take a long time without a good map of the available resources.

In Appendix 8, we provide links to online resources that data users are likely to find the most useful. The resources listed below cover many of the topics discussed in this handbook, but in greater detail:

ACS Data Resources

- FactFinder Help (online help, census data information, glossary, and tutorial)
<<http://factfinder.census.gov/home/en/epss/main.html>>
- 2007 Guide to the Data Products (Web page)
<http://www.census.gov/acs/www/Products/users_guide/>
- American Community Survey Design and Methodology (technical paper)
<<http://www.census.gov/acs/www/Downloads/tp67.pdf>>
- Accuracy of the ACS Data (online document)
<<http://www.census.gov/acs/www/Downloads/ACS/accuracy2007.pdf>>
- ACS Sample Size (Web page)
<<http://www.census.gov/acs/www/SBasics/SSizes/SSizes07.htm>>
- ACS Quality Measures (Web page)
<<http://www.census.gov/acs/www/UseData/sse/>>
- 2006 Data Users Handbook (online document)
<<http://www.census.gov/acs/www/Downloads/Handbook2006.pdf>>
- How to Use the ACS Data (Web page)
<<http://www.census.gov/acs/www/UseData/>>
- Guidance on Comparing 2007 ACS Data (online tables)
<<http://www.census.gov/acs/www/UseData/compACS.htm>>
- Detailed information on ACS Topics
<<http://www.census.gov/acs/www/usedata/Def.htm>>

Conclusion

Now that you've been introduced to the ACS and viewed a small sample of the many ways it can be used in courses from social studies to geography to mathematics and statistics, we hope you are motivated to explore the ACS data on American FactFinder on your own. The ACS provides an unparalleled source of rich information that students and teachers can exploit to learn more about their community and their country while they master concepts and skills in the core subjects of social studies, geography, and mathemat-

ics. It is important to note that this handbook is being written in the summer of 2008, and data from many smaller places are not yet available. However, the first set of ACS products with 3-year estimates will be available by December 2008. Beginning in 2010 when ACS data become available for every geographic level on a regular basis, teachers and students from even the smallest places in the United States will have access to rich information about their communities that is updated each year.

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 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.

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

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.

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.

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 Multi-year 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.

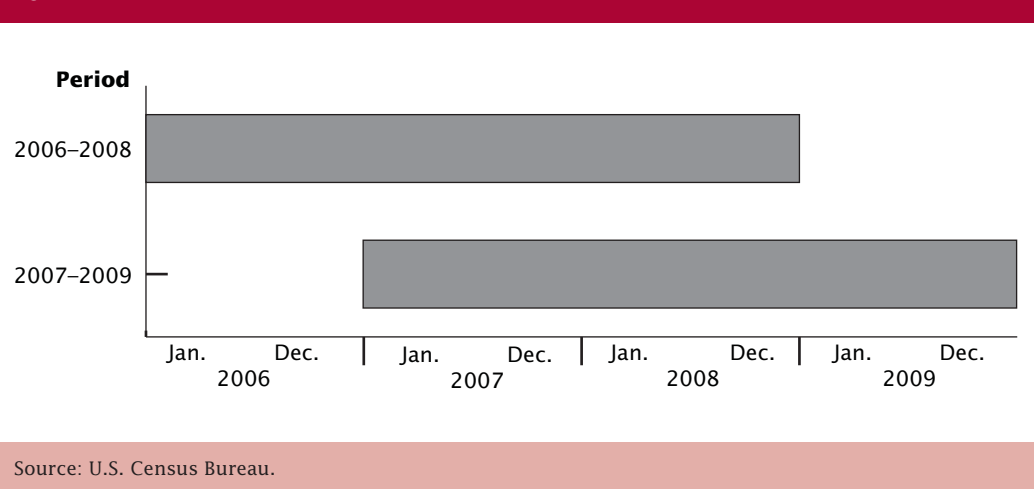
You certainly can compare characteristics for areas with populations of 30,000 to areas with populations of 100,000 but you should use the data set that they have in common. In this example you could use the 3-year or the 5-year estimates because they are available for areas of 30,000 and areas of 100,000.

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

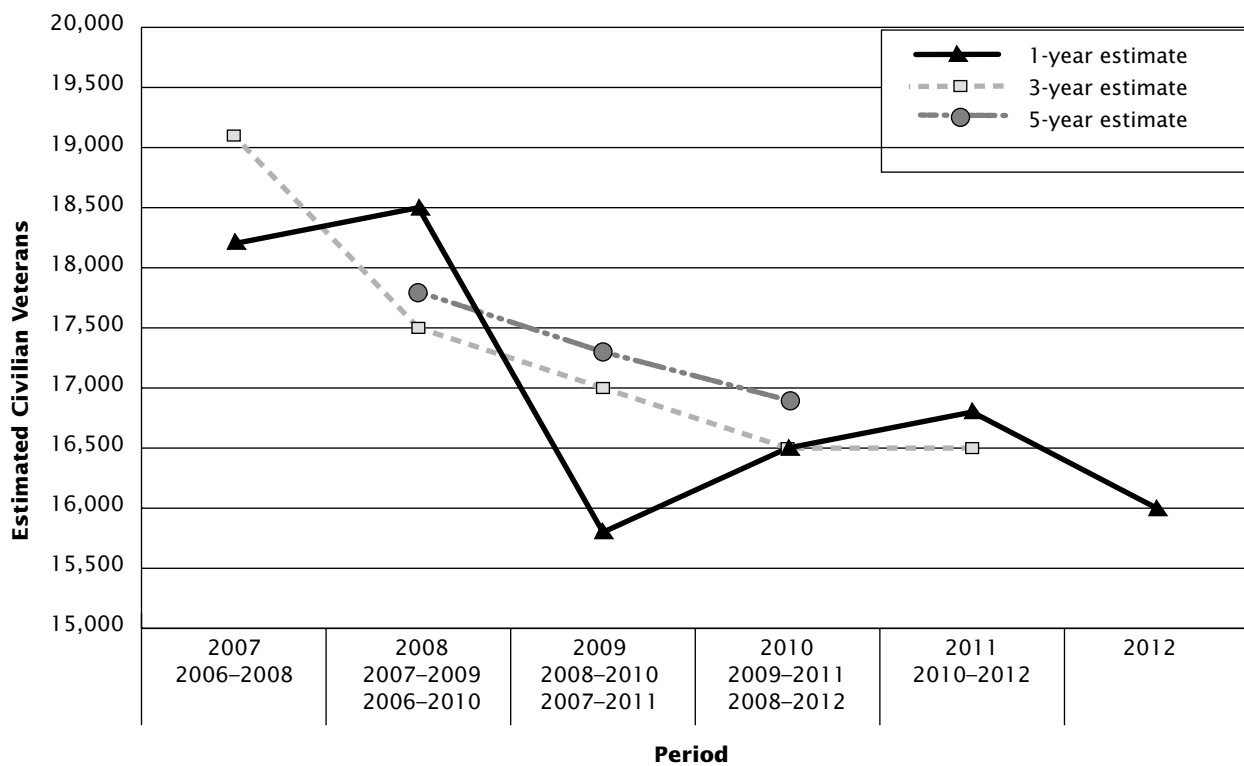


Variability in single-year estimates for smaller areas (near the 65,000-publication threshold) and small sub-groups within even large areas may limit the ability to examine trends. For example, single-year estimates for a characteristic with a high CV vary from year to year because of sampling variation 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 associ-

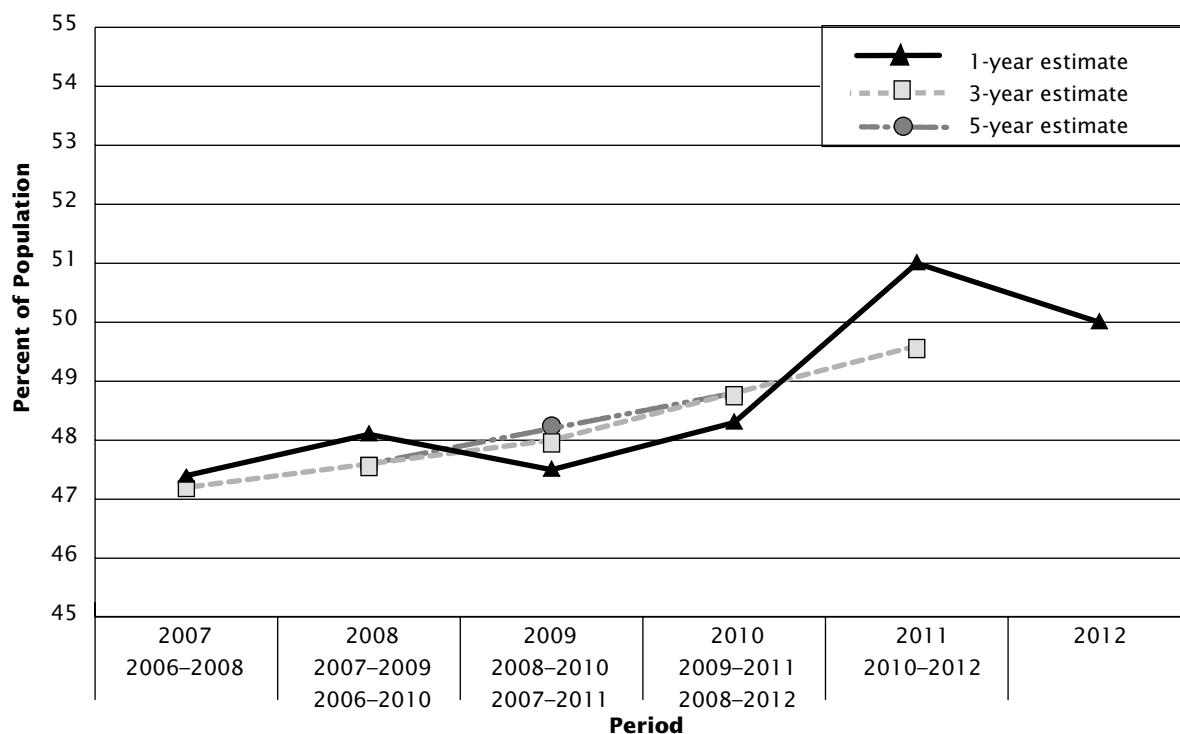
ated with sufficiently small CVs, can be very valuable in identifying and studying such phenomena. Graphing trends for such areas using single-year, 3-year, and 5-year estimates can take advantage of the strengths of each set of estimates while using other estimates to compensate for the limitations 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.

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 $\pm 12,357$. The MOE corresponding to a 95-percent confidence level would be derived as follows:

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

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

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

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.

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 $\pm 12,357$. 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 variabilities. 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 $\pm 1,645$, 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

$$CI_{CL} = (L_{CL}, U_{CL}).^3$$

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 ($\pm 14,723$).

$$L_{95} = 771,782 - 14,723 = 757,059$$

$$U_{95} = 771,782 + 14,723 = 786,505$$

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 ($[1,000 / 20,000] \times 100$). 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)

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

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 | $\pm 3,303$ |
| Females living alone in Arlington County (Component 2) | 19,464 | $\pm 2,011$ |
| Females living alone in Alexandria city (Component 3) | 17,190 | $\pm 1,854$ |

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:

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

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

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

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 $\pm 4,289$.

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.

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

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.

$$\begin{aligned} MOE_{num}^2 + (\hat{R}^2 * MOE_{den}^2) &= \\ (989)^2 + [(0.7200)^2 * (1,328)^2] &= \\ 978,121 + 913,318.1 &= 1,891,259.1 \end{aligned}$$

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 $50,624 \times 0.134 = 6,784$. The MOE is calculated by:

$$\begin{aligned} MOE_{A \times B} &= \pm \sqrt{50,624^2 \times 0.027^2 + 0.134^2 \times 2,423^2} \\ &= \pm 1,405 \end{aligned}$$

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

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

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:

$$\begin{aligned} \text{Percent Change} &= 100\% * \frac{\hat{X}_2 - \hat{X}_1}{\hat{X}_1} = \\ 100\% * \left(\frac{802,210 - 762,475}{762,475} \right) &= 5.21\% \end{aligned}$$

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):

$$\begin{aligned} MOE_2^2 + (\hat{R}^2 * MOE_1^2) &= \\ (22,866)^2 + [(1.0521)^2 * (22,666)^2] &= \\ 1,091,528,529 \end{aligned}$$

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.

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

- 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:

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

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 ($i=1,2$)

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

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:

$$\begin{aligned} \left| \frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}} \right| &= \left| \frac{22.0 - 21.5}{\sqrt{(0.12)^2 + (0.12)^2}} \right| = \\ \left| \frac{0.5}{\sqrt{0.015 + 0.015}} \right| &= \left| \frac{0.5}{\sqrt{0.03}} \right| = \left| \frac{0.5}{0.173} \right| = 2.90 \end{aligned}$$

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.

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

$$\begin{aligned} \left| \frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}} \right| &= \left| \frac{22.5 - 22.1}{\sqrt{(0.20)^2 + (0.25)^2}} \right| = \\ \left| \frac{0.4}{\sqrt{0.04 + 0.0625}} \right| &= \left| \frac{0.4}{\sqrt{0.1025}} \right| = \left| \frac{0.4}{0.320} \right| = 1.25 \end{aligned}$$

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.

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.

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.

