

# Appendix A:

## SOURCE AND ACCURACY OF DATA

### SOURCES OF DATA

The data for this report, which cover a wide range of topics and years, came from the Current Population Survey (CPS), the CPS supplements, the American Community Survey (ACS), the Survey of Income and Program Participation (SIPP), the American Housing Survey (AHS), and the U.S. Census Bureau's Population Estimates Program. The CPS supplements used for this report include the Annual Social and Economic Supplement (ASEC) and supplements on child support, computer and Internet use, school enrollment, fertility, and voting and registration.

This report includes data for four different population universes: the resident population (census universe); the civilian noninstitutionalized population (CPS supplements universe); the civilian noninstitutionalized population, plus armed forces living off post or with their families on post (SIPP and CPS ASEC universe); and the universe of housing units (AHS). The 2005 data for the ACS are limited to the household population and exclude the population living in institutions, college dormitories, and other group quarters. Beginning with the 2006 data, the ACS universe will be the resident population.

Estimates using sample data from the CPS for 2001 and earlier and from the SIPP are weighted by population controls. Controls are based on updated 1990 decennial census data adjusted for an estimated net undercount. As such, these estimates are not consistent with population estimates computed from the intercensal estimates program, which are not adjusted for an estimated net census undercount. Data from the CPS for 2002 or later and the SIPP 2001 Panel or later are weighted using controls based on Census 2000. The ACS uses population and housing controls based on Census 2000. The AHS uses housing unit estimates from Census 2000. ACS and AHS estimates may not be consistent with population estimates from the intercensal estimates program.

For additional information about the CPS and CPS supplements, see <[www.bls.census.gov/cps/mdocmain.htm](http://www.bls.census.gov/cps/mdocmain.htm)> and <[www.census.gov/hhes/www/childsupport/childsupport.html](http://www.census.gov/hhes/www/childsupport/childsupport.html)>. For SIPP, see <[www.sipp.census.gov/sipp/sourceac/S&A96\\_030228.Long.pdf](http://www.sipp.census.gov/sipp/sourceac/S&A96_030228.Long.pdf)>, <[www.sipp.census.gov/sipp/sourceac/s&a96\\_040501.pdf](http://www.sipp.census.gov/sipp/sourceac/s&a96_040501.pdf)>, and <[www.sipp.census.gov/sipp/usrguide/sipp2001.pdf](http://www.sipp.census.gov/sipp/usrguide/sipp2001.pdf)>. For ACS, see <[www.census.gov/acs/www/UseData/Accuracy/Accuracy1.htm](http://www.census.gov/acs/www/UseData/Accuracy/Accuracy1.htm)>. For AHS, see <[www.census.gov/hhes/www/housing/ahs/ahs01/appendixb.pdf](http://www.census.gov/hhes/www/housing/ahs/ahs01/appendixb.pdf)> and <[www.census.gov/hhes/www/housing/ahs/ahs01/appendixd.pdf](http://www.census.gov/hhes/www/housing/ahs/ahs01/appendixd.pdf)>. See also the sources listed at the end of every chapter.

The Population Estimates Program publishes total population estimates each year. The publication of population estimates also includes demographic components of change (births, deaths, and migration). The estimates are also published by age, sex, race, and Hispanic origin. For further information about the Population Estimates Program, see Population Estimates Methodology at <[www.census.gov/popest/topics/methodology](http://www.census.gov/popest/topics/methodology)>.

### ACCURACY OF THE ESTIMATES

A sample survey estimate is subject to two types of error: sampling and nonsampling. The accuracy of an estimate depends on the extent of both types of error. The nature of the sampling error is known given the survey design; however, the full extent of the nonsampling error is unknown.

#### Sampling Error

Since CPS, ACS, SIPP, and AHS estimates come from samples, they may differ from measures from an enumeration of the entire population using the same questionnaires, instructions, and interviewers. For a given estimator, the difference between an estimate based on a sample and the measure that would result

if the sample were to include the entire population is known as sampling error.

Standard errors are primarily measures of the magnitude of sampling error. They are not given in this report because of the wide range of topics included and the wide variety of data sources. Standard errors may be found in the publications that are noted at the end of most sections of the Population Profile.

Because the estimates in this report (which may be shown in text, figures, and tables) are based on responses from a sample of the population and may differ from actual values because of sampling variability or other factors, apparent differences between the estimates for two or more groups may not be statistically significant. All comparative statements have undergone statistical testing and are significant at the 90-percent confidence level unless otherwise noted.

### **Nonsampling Error**

For a given estimator, the difference between the estimate that would result if the sample were to include the entire population and the true population value being estimated is known as nonsampling error.

To minimize these errors, the Census Bureau employs quality control procedures in sample selection, wording of questions, interviewing, coding, data processing, and data analysis.

### **Comparability of Data**

Data obtained from sample surveys and other sources are not entirely comparable due to differences in interviewer training and experience, different survey processes, and differences in the target population.

These examples are sources of nonsampling variability not reflected in the standard errors. Therefore, caution should be used in comparing results from different sources.

Caution should be used when comparing data from a microdata file that reflect Census 2000- based population controls with data from microdata files from March 1994–December 2001, which reflect 1990 census-based population controls. Caution should also be used when comparing the data from a microdata file that reflect 1990 census-based population controls with data from microdata files from March 1993 and earlier years, which reflect 1980 census-based population controls. When comparing data within microdata files, be sure to use estimates that reflect the same population controls. Microdata files from previous years reflect the census-based population controls for the estimate date that were most current when the estimates were made. Although this change in population controls had relatively little impact on summary measures such as averages, medians, and percentage distributions, it did have a statistically significant impact on levels. For example, use of Census 2000-based population controls results in about a 1 percent increase from the 1990-based population controls in the civilian noninstitutionalized population and in the number of families and households. Therefore, estimates of levels for data collected in 2002 and later years will differ from those for earlier years by more than what could be attributed to actual changes in the population. These differences could be disproportionately higher for certain population groups than for the total population.