Meeting 21st Century Demographic Data Needs—Implementing the American Community Survey: May 2002

Report 2: Demonstrating Survey Quality

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

Acknowledgments

U.S. Department of Commerce
Donald L. Evans,
Secretary

ECONOMICS AND STATISTICS ADMINISTRATION

Economics and Statistics Administration
Kathleen B. Cooper,
Under Secretary for Economic Affairs

U.S. CENSUS BUREAU

Charles Louis Kincannon,
Director

William G. Barron, Jr.,
Deputy Director

John H. Thompson,
Principal Associate Director for Programs

Deborah H. Griffin and Sally M. Obenski were the key authors of this report.

Members of the American Community Survey Research and Evaluation Steering Committee who sponsored and reviewed this report include:

Alan R. Tupek, Chair
Cynthia Z. F. Clark
Nancy M. Gordon
Preston J. Waite
Chester E. Bowie
Ruth Ann Killion
John F. Long
Elizabeth A. Martin
Daniel H. Weinberg
Tommy Wright
Deborah H. Griffin
Sally M. Obenski

Evaluation Project Managers whose data are reflected in the report include: Keith A. Albright, Charles H. Alexander, Deborah H. Griffin, Susan P. Love, Alfredo Navarro, Michael D. Starsinic, and Anthony G. Tersine.

Other individuals who contributed to the review and release of this report include: Florence H. Abramson, Lisa E. Buckley, Jennifer A. Guarino, David L. Hubble, Juanita T. Lott, Elizabeth A. Martin, and Virginia M. Zamaitis.

Table of Contents

Appendices

Tables and Figures

Executive Summary

As part of the American Community Survey (ACS) development program, the Census Bureau initiated a research and testing program to examine a range of ACS implementation issues. Key results are being documented in a series of reports. This is the second report. It briefly discusses timeliness and focuses on the accuracy of the Census 2000 Supplementary Survey (C2SS). Future reports will further assess C2SS data quality by comparing responses to questions on the decennial census short and long forms.

The impetus for the development of the ACS was to greatly improve the timeliness of detailed socioeconomic and demographic—decennial census long form—data while maintaining quality. Over the decade, Census 2000 long form data products will continue to age as ACS data products begin providing an ongoing profile of the nation's people and economy. However, replacing the decennial census long form implies that the ACS must achieve an acceptable level of quality. The expected outcome is that while sampling error will slightly increase, nonsampling error should be decreased when compared to the decennial census long form. Although not exhaustive, this report includes analyses of error levels in the C2SS, and how such error affects the survey's performance.

Accuracy measures indicate that the C2SS was of high quality, indicating that the ACS will have the quality to replace the decennial census long form, beginning in 2003. Accuracy refers to the closeness between estimated and true (unknown) values. High levels of survey error can lead to incorrect conclusions. Both sampling and nonsampling error can affect accuracy. Sampling error occurs because the survey is using a proportion or sample of the population upon which to draw its inferences. Sampling error is mathematically derived and is, therefore, straightforward to calculate. Nonsampling error includes all other errors including nonresponse, coverage measurement, and processing errors. Nonsampling error is far more difficult to detect and to measure. While the errors herein are depicted as distinct, they are, in fact, overlapping. In developing standards for a large-scale survey like the ACS, designers looked to premier surveys and the decennial census. Where appropriate and when data were available, comparisons were made to the decennial census long form. Major observations include:

Traditional measures suggest nonresponse error is under control (see page 15). Survey nonresponse is the most visible and well-known source of nonsampling error and has two types—unit and item nonresponse. Unit nonresponse is the failure to obtain data from a unit in sample for the survey. Item nonresponse occurs when a responding unit fails to provide complete and usable information for data items, such as income. Because data were not readily available, this report includes limited comparisons to Census 2000. The results of such comparisons are the subject of the next reports in the series.

> **Unit Nonresponse** (see page 15). The C2SS survey response rate was 95.4 percent and includes responses from the combination of mail, Computer Assisted Telephone Interviewing (CATI), and Computer Assisted Personal Interviewing (CAPI). Although impressive, closer examination of survey response rates for areas with high concentrations

of racial and ethnic population groups indicated that some groups experienced lower survey response rates than those for the White population. For example, predominantly White areas had a survey response rate of 95.9 percent while areas predominantly African American or Black and American Indian or Alaska Native had survey response rates of 92.4 percent and 89.1 percent, respectively.

Item Nonresponse (see page 19). As in most surveys, the C2SS uses imputation methods to handle item nonreponse. Values may be assigned based on other information on the questionnaire, or statistically derived based on the characteristics of respondents thought to be similar to the nonrespondent. The basic population items (i.e., relationship, gender, age, Hispanic Origin, and race) contained in Census 2000 short and long forms were examined**.** When the C2SS imputation rates were compared to Census 2000, C2SS rates were significantly lower for every basic population item.

Sample completeness ensured accurate survey representation (see page 22). The C2SS sample completeness, which measures the extent to which the interviewed survey sample represents the Census 2000 population count, was about 90 percent—slightly higher than the sample completeness for the 1990 decennial census long form survey. The closer the measure is to 100 percent, the higher the completeness level. Several factors can account for the ten percent difference between the C2SS and Census 2000. One important factor was that the C2SS and Census 2000 used different versions of the Master Address File as sampling frames. Other factors include unit nonresponse (discussed above) and survey undercoverage. Survey undercoverage results from housing units or household members not being included in the survey when they should have been. Survey undercoverage refers to a very different aspect of coverage than census undercoverage. As measured as a comparison to the Census 2000 population count, the C2SS sample completeness measure does not reflect any housing units or persons missed in both Census 2000 and the survey.

Processes exist to control measurement and processing error (see page 25). Procedures have been put into place to help control measurement and processing errors in the C2SS. Such procedures have the potential to result in lower measurement and processing error levels than are found in the decennial census long form survey**.** Measurement error refers to the difference between the observed value of a variable gathered during data collection and the true, unobserved value of the variable. Processing error occurs during the series of operations that convert reported data into consistent, machine-readable information and published estimates. Procedures to control measurement and processing error are either inherent to the use of the ACS methods or are applied as part of C2SS quality assurance. Quality assurance activities included a formal quality control reinterview as part of the CAPI operation and a detailed program for managing the level of errors introduced during data entry.

Sampling error was controlled and will be slightly higher than the decennial census long form survey (see page 27). Sampling error levels in the ACS test sites met survey objectives. These levels will be somewhat higher for the ACS than for the decennial census long form survey. Sampling error arises from the use of a sample to represent a population. The larger the sample is that represents the population, the smaller the sampling error. Therefore, sampling error

increases anytime the sample is reduced. One measure of sampling error—the Coefficient of Variation or CV—attempts to quantify the relationship between the size of the error and the size of the estimate. The smaller the CV, the more precise the estimate. In general, the stated target was that the CVs would typically be about 12 percent for single-year ACS estimates for areas of 65,000 population. In fact, the actual mean and median CVs of interest were computed as 12.2 percent and 11.8 percent, respectively**.** Based on these results, it is anticipated that CVs for small areas will be slightly higher than the CVs for the decennial census long form survey.

Overview and Purpose

Census Bureau managers have concluded that decennial census operational complexity must be decreased; and data users, as well, have long argued that the currency of detailed housing and population data should be increased. The 2010 reengineering strategy will manage risk and reduce complexity while improving decennial census coverage and containing costs in the 2010 census. The American Community Survey (ACS) is one of three program components required to achieve the 2010 reengineering strategy goals. Collecting long form data throughout the decade by the ACS will have a profound effect on the design, planning, and potential outcome of the 2010 census.

Over 10 years ago, in response to congressional and other stakeholder demands for timely and relevant data, the Census Bureau began examining a new approach for gathering long form data. Consequently, in lieu of the static, once-a-decade snapshot of the nation's population, Census Bureau experts began researching the feasibility of an ongoing survey to collect and disseminate timely demographic and socioeconomic data. This research culminated in 1994 with the initiation of the ACS development program. Since then, the program's name, size, and scope have been evolving in preparation for implementation in 2003.

The primary purpose of the ACS development program was to develop the methods for providing timely, accurate, and detailed long form data each year. Over a period of 4 years, this activity expanded from 4 test sites to 36 counties contained in 31 ACS test sites. The Census 2000 Supplementary Survey (C2SS) was conducted as part of Census 2000 in 1,203 additional counties using ACS methods. Its primary purpose was to demonstrate the operational feasibility of collecting long form data at the same time as, but separate from, Census 2000. The C2SS, combined with the 36 counties contained in the ACS test sites, provided national level data. Data collection activities for the 2001 and 2002 Supplemental Surveys have been continuing in the same 1,203 counties. These surveys will allow multi-year estimates to be produced and will help demonstrate the data's usability and reliability. Operational testing activities will be completed by the end of 2002, culminating in the nationwide implementation of the ACS.

When the term "ACS development program" is used herein, it refers to the full set of testing, research, and development program activities that started in 1994 and will continue until the ACS is fully implemented in fiscal year 2003. A more detailed description of the ACS development program activities can be found in Appendix VI, the Glossary of Terms. While major operational testing activities will be completed upon implementation, a full suite of ACS improvement and enhancement activities will be ongoing throughout the life of the survey.

As part of the Census Bureau's comprehensive ACS development program, key results will be documented in a series of reports. This is the second report in the series and focuses on measures of C2SS survey quality. Specifically, this report summarizes sampling and nonsampling error levels in the C2SS. This report is not intended to include an exhaustive set of quality measures but those that are generally accepted as key measures that are currently available. This report does not provide detailed analysis on the quality of the estimates produced from the C2SS.

Subsequent reports, comparing estimates from the C2SS and Census 2000, will include the full spectrum of items contained in both the decennial census short and long form questionnaires.

In preparing this report, Census Bureau analysts consulted with survey experts to identify key technical quality indicators, analyzed operational and evaluation data, and documented the results. Further, statistical methods publications were consulted to ensure consistency with generally accepted survey quality measures. Due to its size, complexity, and role as the replacement for the decennial census long form, quality indicators from both survey and census literature were incorporated.

Analysts reviewed available information from both the C2SS and the 36 counties contained in the ACS test sites. Almost all of the results presented in this report are based on the combination of these data. For ease of documentation, however, the combined data will be referred to as the C2SS.

Over the next year and a half, the Census Bureau will release the following additional reports:

- A detailed analysis of basic demographic characteristics (e.g., relationship, race, tenure) produced from the C2SS at the national and state levels, including comparisons between the C2SS and Census 2000;

- A description of the data release plan and products for the ACS and the usability and accessibility of estimates resulting from ACS methods; and,

- Several detailed analyses of selected social, economic, and housing characteristics (e.g., education, income, commuting patterns), including, comparisons between the C2SS and Census 2000 at the national and some sub-national levels.

Introduction

Charged by the Congressional Act of March 6, 1902, to be the nation's data collector, the Census Bureau is the only federal agency that has the infrastructure and capacity to conduct multiple national, large-scale household surveys and censuses. The Census Bureau provides the Administration; the Congress; the business community; state, local, and tribal planners; trade associations; academicians; and other data users with a vast array of essential information. The Census Bureau also conducts *periodic* censuses—most notably, the decennial census of population and housing.

For decades, decennial census long form data have played an indispensable role in governing the nation. Every question on the form has a specific federal legislative, regulatory, or judicial mandate. The Census Bureau reported to the Congress on March 30, 1998, on Census 2000 questions, providing selected citations to about 130 laws that use specific long form data items. As one example, 25 citations alone exist for the disability questions on the long form. Without decennial long form data, the government would lack the information required by law to effectively manage its programs.

As a result of the decennial census long form, the Census Bureau has been delivering timely, relevant, high quality demographic and socioeconomic data to its stakeholders. Rapid demographic and technological change, however, are outpacing the Census Bureau's ability to deliver such data through a decennial collection activity. This fact has prompted managers to adopt a strategy that harnesses the synergy among three Census Bureau programs—the MAF/TIGER[1] modernization, the 2010 decennial census planning and testing, and the ACS implementation. Beginning in 2004, the ACS will provide continuous, detailed demographic and socioeconomic data, replacing the once-a-decade decennial census long form survey. Consequently, the Census Bureau must demonstrate that the ACS is a quality survey that delivers high quality data that will meet 21st Century demographic data needs.

Four dimensions comprise survey quality

It is important that both the producers and consumers of survey data become knowledgeable about the elements of survey quality. According to the Federal Committee on Statistical Methodology, survey quality has four key dimensions: relevance, accessibility, timeliness, and accuracy.[2]

This report focuses on timeliness and accuracy. Timeliness refers to both the length of time between the data collection and the first availability of a product and to the frequency of the data collection. Data are considered timely if a minimal amount of time separates the event described by the data and the data availability. Data are current if the time frame for data collection is close to the time frame for data use. Data from recurring surveys, such as the ACS, produce current

[1] Master Address File/Topologically Integrated Geographic Encoding and Referencing System.
[2] "Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001.

data while periodic or one-time survey data may quickly become obsolete. Accuracy refers to the closeness between estimated and true (unknown) values. This is probably the most important aspect of survey quality. Accuracy is measured via sources of survey errors. High levels of survey error can lead to incorrect conclusions by data users.

Subsequent content and data product reports in the series will include the dimensions of relevance and accessibility. Relevance is a qualitative measure of data value. Any assessment of relevance must acknowledge the major uses of the data (e.g., legislative) and evaluate how well the data meet these needs. The survey's ability to collect the most meaningful data needs to be evaluated. Studies of relevance must also consider if specific uses of survey data fit within the general purposes for which the survey was designed. Accessibility refers to the ability of the data users to readily obtain and use survey products. It encompasses dissemination media, availability of metadata, ease of access, and affordability.

Although survey errors can be summarized in different ways, they fall into two broad categories—sampling and nonsampling errors. Sampling error refers to the variability that occurs by chance because a sample, rather than an entire population, was surveyed. Nonsampling error refers to all other errors that occur in a survey, such as nonresponse (missing or incomplete information from the sample), coverage (missing or duplicate units or persons), measurement (data collection errors), and processing errors.

Generally speaking, survey designers must make trade-offs not only between sampling and nonsampling error but also among the other three dimensions of survey quality—relevance, accessibility, and timeliness. The trade-offs are determined through analyses of cost, schedule, and required performance. For example, the decennial long form has been the standard for small area socioeconomic data. However, as mentioned above, concerns about the timeliness of the decennial long form data have prompted managers to adopt the ACS as a replacement.

Census Bureau experts and managers concluded that, when implemented, the ACS will improve overall survey quality compared to the decennial census long form. That is, some increase in sampling error will occur due to smaller sample sizes in any given year. However, timeliness will greatly improve, and nonsampling error should be reduced by the use of permanent, highly trained field staff.

ACS methods provide a reasonable replacement for the decennial long form survey

ACS methods are designed to ensure the continuing collection and dissemination of detailed demographic and socioeconomic data previously collected during the decennial census. The survey's sample design, data collection and processing operations, and adjustments, known as weighting techniques, work together to help deliver ongoing, high quality data. In contrast, the design of the decennial census long form survey is constrained by the need to give the highest priority to producing high quality population counts in a short timeframe. Consequently, it is not feasible to optimize methods for collecting detailed demographic and socioeconomic data.

Further, once designed and implemented, decennial census long form survey data collection methods cannot be readily modified.[3] Because the ACS is ongoing, opportunities exist for continual improvement as the decade progresses.

Despite smaller sample sizes, ACS estimates will be reasonable replacements for decennial census long form estimates, including those for small areas

The decennial census long form is a survey of about 17 percent of the population. This size sample is required to produce reliable estimates for small areas (e.g., census tract) and small population groups (e.g., American Indian). By comparison, the ACS will occur annually and sample about 2.5 percent of the population each year. The ACS will provide estimates of demographic, housing, social, and economic characteristics every year for all states, as well as for all cities, counties, metropolitan areas, and population groups of 65,000 people or more.

Areas of 20,000 to 65,000 can use data averaged over 3 years. For rural areas and city neighborhoods or population groups of less than 20,000 people, it will take 5 years to accumulate a sample that is sufficient to produce the same estimates as the decennial census. These multi-year averages can be refreshed every year, keeping them up-to-date. Eventually, change can be measured over time for small areas and population groups.

ACS methods employ a comprehensive set of procedures

The ACS development program data are collected in continuous, 3-month cycles using a combination of mailout/mailback, Computer Assisted Telephone Interviewing (CATI), and Computer Assisted Personal Interviewing (CAPI) data collection modes. The ACS design relies on optimizing these three modes of data collection. A sample of housing unit addresses is drawn from the MAF, and questionnaires are mailed out at the beginning of month one. About 3 weeks later, a second questionnaire is mailed to those who did not respond by mail. Telephone Questionnaire Assistance (TQA) is available to help households complete the forms that they receive in the mail. Forms returned by mail are keyed and the data file is reviewed for completeness and consistency. Incomplete forms are included in the Edit Followup operation that contacts respondents by telephone to obtain the missing information.

During month two, interviewers conduct a CATI operation. The interviewers contact housing units from which a mail response has not been received and for which telephone numbers have been obtained from vendors. Finally, in month three, a subsample of the remaining nonrespondents is drawn and professional interviewers conduct personal visits to those housing units. The use of a computerized questionnaire, rather than paper, allows the Census Bureau to incorporate data consistency checks into the data collection process.

[3] "Meeting 21st Century Demographic Needs—Implementing the American Community Survey: Demonstrating Operational Feasibility," U.S. Census Bureau, July 2001.

Figure 1 ACS Primary Process Flow Diagram

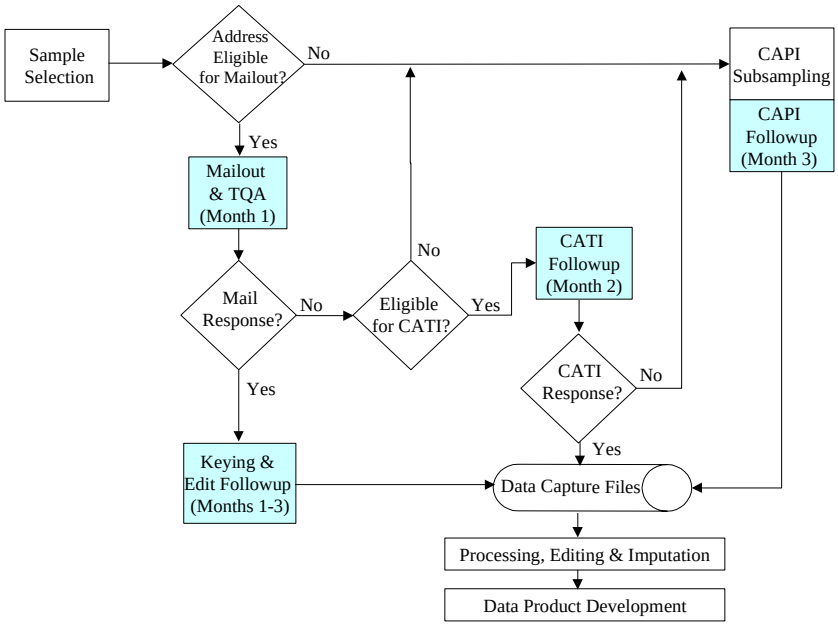

Figure 1 depicts the ACS development program's operations. The first ACS report on operational feasibility provided important details about each of these data collection and processing activities.[4] To better understand how this multi-mode design could impact survey quality, each data collection mode—mailout/mailback, CATI, and CAPI—was evaluated for operational efficiency in the first ACS report. Real-time measures of success in obtaining responses from eligible units by mail, CATI, and CAPI are calculated. These operational response rates are monitored and reported on a regular basis.

These rates serve two important functions—to identify possible performance problems and to shed light on potential nonresponse error by characterizing nonresponse at each stage of data collection. Response rate definitions follow guidelines from Statistical Policy Working Paper 31 and the American Association of Public Opinion Research (AAPOR), which recommend that response be measured relative to the universe of eligible cases.[5] [6] Detailed definitions are in Appendix I. Table 1 includes the mail, CATI, and CAPI operational response rates for the C2SS data collection activities taking place between January and December 2000.

[4] "Meeting 21st Century Demographic Needs— Implementing the American Community Survey: Demonstrating Operational Feasibility," U.S. Census Bureau, July 2001.
[5] "Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001.
[6] American Association for Public Opinion Research. 2000. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys," Ann Arbor, MI: AAPOR.

Table 1: The C2SS Operational Response Rates

Data Collection Mode	Operational Response Rate (in percent)
Mail	51.9
CATI	57.3
CAPI	91.7

The mail response rate is the ratio of all completed mail returns relative to the mailout universe. As shown in Figure 1, some sample addresses are determined not eligible for mailing. About 4.5 percent of the C2SS sample addresses were not eligible for the mailout and were excluded from the denominator of the mail response rate. The C2SS mail response rate was 51.9 percent. Census 2000 estimates that the mail response rate for the long form was 53.9 percent as of the start of nonresponse followup activities.[7] Although the decennial census uses slightly different definitions to calculate these rates, they are provided as a benchmark for the C2SS rates. Studies are planned to review mail response rates and better understand where special efforts might be needed to improve the efficiency of mailout/mailback methods for the ACS.

The CATI response rate is the ratio of completed CATI interviews to the total cases eligible for CATI interviewing. Sample addresses that do not respond by mail and have a valid phone number are considered eligible. Figure 1 indicates that sample addresses not eligible for CATI and noninterviews from CATI are eligible for CAPI subsampling. The CATI response rate for the C2SS was 57.3 percent. This rate provides a measure of the success of collecting the C2SS data by telephone—once a valid phone number was obtained. CATI noninterviews could have been due to refusals, language barriers, callbacks that were never converted to completed responses, or the inability to contact a household member. Analysis is needed to determine the major reasons for CATI nonresponse. The overall effectiveness of the CATI operation can be improved if better sources can be found to obtain valid telephone numbers.

The CAPI response rate is the ratio of completed CAPI interviews for both occupied and vacant units to the total cases eligible for CAPI interviewing. This eligible universe excludes addresses determined to be commercial or nonexistent. The universe includes addresses that were not eligible for mail and CATI interviewing and noninterviews from these two data collection efforts. The CAPI response rate for the C2SS was 91.7 percent. Noninterviews could have been due to refusals, noncontacts, language barriers, or unresolved callbacks to nonresponding housing units. As with CATI, detailed analyses are needed of the reasons for nonrespone. The geographic distribution of these CAPI noninterviews is currently being studied. The research will determine if CAPI nonresponse is clustered.

Weighting techniques calibrate survey results to improve overall quality

C2SS estimation includes a series of weighting adjustments. Survey data are traditionally weighted to adjust for the sample design, the effects of nonresponse, and to correct for survey

[7] Stackhouse, Herbert F. and Treat, James B., "Census 2000 Response and Return Rates—National and State by Form Type," DSSD Census 2000 Procedures and Operations Memorandum Series L-10, February 12, 2002.

undercoverage. ACS developers are currently reviewing all of these weighting techniques to see if changes are warranted. Initial weights are applied to the C2SS sample units to account for differences in selection probability resulting from the sample design. For example, each unit, sampled at a rate of 1 in 40, gets a weight of 40. In oversampled small governmental units where the rate is 1 in 10, each unit gets a weight of 10. When units that have not responded by mail or CATI are subsampled for CAPI at a rate of 1 in 3, their weight is multiplied by 3.

A second weighting adjustment is for unit nonresponse—when a household identified for interview does not respond or so little data are obtained that they cannot be used to produce estimates. In the C2SS, higher weight is given to interviewed units in a given tract and month to account for noninterviews in that tract and month. For example, if only nine-out-of-ten of the designated units are interviewed in a tract in a specific month, a nonresponse adjustment of 10/9 is used to increase the weight of the interviewed units when they are included in the estimates.

A final weight is applied to ensure that the survey results are adjusted for survey undercoverage. That is, the adjustment helps to ensure that the distributions of the characteristics being collected (e.g., age, race, sex) are comparable to the standard—the decennial census or the intercensal estimates that are based on the decennial census. This so-called "weighting to population control totals" also compensates for errors not corrected by the other two weighting techniques.

ACS methods support continued design improvement

While ACS methods are robust, they can be further optimized as additional data become available and are analyzed. Consequently, improving the quality of ACS methods is an ongoing process. In addition to specific analyses (e.g., examining the weighting techniques), evaluations of operational and quality assurance activities provide performance information so that corrective actions can be taken. For example, the initial ACS report that demonstrated operational feasibility revealed that a substantial percentage of cases eligible for Edit Followup were not processed due to a staffing shortfall.[8] Subsequently, staffing levels were increased. Additionally, after close review of how survey response rates have been calculated, an adjustment has since been made to more completely account for all types of noninterviews. Assessing, analyzing, and initiating needed adjustments will continue to require dedicated resources.

While the initial ACS report demonstrated the operational feasibility of implementation of the ACS, it also noted that completing the survey on time, with available resources, and within budget does not by itself indicate a quality survey.[9] This report provides some insight into the quality of the C2SS, beginning with a brief discussion of timeliness.

[8] "Meeting 21st Century Demographic Data Needs—Implementing the American Community Survey: Demonstrating Operational Feasibility," U.S. Census Bureau, July 2001.
[9] Ibid.

Timeliness of ACS Data is an Advantage over the Long Form

To date, most empirical data on the usefulness of the ACS development program are found at the local level, which is appropriate. The ACS or long form content is aimed at topics that are needed for geographic units smaller than the nation and states, because the data are federally mandated at the census tract level or for very small population groups. If information were needed only at the national or state level, the ACS or decennial long form survey would not be required since other surveys can provide such data. Beginning in 2003, the ACS will collect detailed demographic and socioeconomic data. By 2008, four to five years before decennial census long form data would have been available, the ACS will be providing detailed profiles for the smallest areas and groups.

Timely availability of ACS data products enhances data currency over decennial long form products

As the decade progresses, census long form data products become increasingly outdated, while ACS data products will continue to provide timely and useful detailed demographic and socioeconomic data. Census long form data products can be as old as 12.5 years before being replaced with the next set of census long form data. In contrast, the ACS will be updating data products yearly for communities. The oldest an ACS data product can be, even for the smallest communities, is 3.5 years. For comparison, census long form data at their most current age are between 2 and 2.5 years old.

The "age" of a data product refers to the length of time between the reference date for the data collection and the date when the product is being used. For example, the reference date for the 1990 census long form was April 1, 1990, and the last of the Census 2000 Summary File 3 data products are scheduled for release at the end of September 2002. So the main 1990 census long form data products are 12.5 years old when they are fully replaced.

For the ACS, the reference date refers to an entire year, or several entire years. The mid-point of the year is used as the reference point to represent the year. So the 2011 annual estimates have a reference date of July 1, 2011. The 5-year estimates (covering 2007 – 2011) have a reference date of July 1, 2009 – the mid-point of the 5-year period. Therefore, in the period of July through December of 2012, when these estimates would be released, the annual data (2011) would be 1 to 1.5 years old, and the 5-year average data (2007 – 2011) would be 3 to 3.5 years old. If there were a census long form in 2010, its data would likely be released in this same time period and would be 2 to 2.5 years old when released.

The ACS will produce information on content items similar to the decennial census long form for communities in the country, including small areas such as census tracts, small towns, American Indian Reservations, Native Alaskan villages, and rural areas. It will take 5 years for sufficient sample to be collected by the ACS to produce reliable estimates for such small areas and groups. However, once the small area data products (i.e., 5-year data products) are produced in 2008, from that point on, they will be updated and available yearly.

By 2008, Census 2000 long form data products will be between 8 and 8.5 years old. For the smallest areas, they will be replaced by the ACS data products that are only 3 to 3.5 years old and will be updated annually thereafter. If there were a Census 2010 long form, it would not replace the Census 2000 data products until 2012, with data products that would be 2 to 2.5 years old and not updated thereafter.

Although the small area data is of critical importance, as an added benefit, the ACS will also be producing many other data products for larger geographic areas. The first data for communities of 65,000 people or more (i.e., single-year data products) will be available in 2004, about six months after the 2003 data collection year ends and every year thereafter. The first data for communities with between 20,000 and 65,000 people or more (i.e., 3-year data products) will be available in 2006 and updated yearly.

By 2008, the ACS will have profiled the entire nation at all size geographical levels and groups. As discussed above, once profiled, from this point forward, all products will be refreshed on a yearly basis. The ACS will allow for the identification of key changes and will provide much more meaningful data for areas with marked changes in local demographics and economics since the decennial. As indispensable as the decennial census long form survey has been to governance, as seen below, a once-a-decade snapshot is no longer adequate for capturing the dynamics of the 21st century.

The timeliness of ACS data can support strategic planning and disaster recovery

The usefulness of the information is indicative of the value that the fully implemented ACS will bring to federal and state governments. As discussed in the first report, federal, state, and local officials have already been considering how to take advantage of the timely ACS development program data.[10] The most poignant use of the ACS development data is in assisting government officials in disaster planning and recovery in the aftermath of the September 11, 2001 attacks on New York City and on the Pentagon. While one ACS test county is assessing how to use the ACS development data for emergency preparedness against a future attack, New York City will need the data to assist in recovering from the attack that has already occurred. Because of the magnitude of the destruction in New York City, the detailed socioeconomic data collected in Census 2000 in communities surrounding the Trade Centers will describe the earlier—pre-attack—New York City. Without the ACS development program, New York officials would have to wait until 2012 before fully understanding the demographic and socioeconomic changes that have occurred since the attack.

In contrast, the C2SS will provide a picture of New York City before the attacks that can be compared to data collected in the 2002 Supplementary Survey. Comparing the two sets of detailed demographic and socioeconomic data will provide officials with a comprehensive picture of the magnitude and direction of the effects. For example, the comparison may provide insight

[10] Ibid.

into the effect on average income in those communities most affected. City planners can use timely, relevant ACS development program data on occupation patterns or travel to work and other factors to assist in and track recovery.

Although timeliness is an important dimension of quality and essential for an informed recovery for New York City, if the data are not accurate, the value of timeliness is diminished. The remainder of this report describes and measures key components of the C2SS accuracy.

Traditional Measures Suggest Nonresponse Error is Under Control

Nonresponse is the most visible and well-known source of nonsampling error. There are two main types of nonresponse error – unit nonresponse and item nonresponse. Nonresponse errors exist in surveys and affect survey estimates to varying degrees. The exact amount of nonresponse error on an estimate is almost never known. Nonetheless, proxies for nonresponse error, which measure specific aspects of survey quality, are critical to informing data users of the usefulness and limitations of the data. This report relies primarily on survey response rates to assess unit nonresponse error. Imputation rates were calculated to measure error introduced by item nonresponse. Survey response and item imputation rates do not provide complete measures of nonresponse error. The amount of nonresponse error introduced is a function of both the response and imputation rates and how different the characteristics of nonrespondents are from respondents.

Controlling unit nonresponse maintains high survey quality

Unit nonresponse is the failure to obtain data from a unit in the sample and may occur because households are unwilling or unable to participate, or because an interviewer may be unable to make contact with a respondent for a sample unit. Unit nonresponse results in estimates that represent a population short of the true population. If the demographic, social, economic, or housing characteristics of noninterviewed households differ from the households that are interviewed, the ACS will not provide an accurate picture of the whole population. Maintaining low levels of unit nonresponse and understanding the causes and effects of the errors are important. Measures for unit nonresponse in the Census 2000 long form survey were not available at the time of this report, but will be included in the next reports in the series.

High survey response rates suggest minimal error introduced by unit nonresponse

Survey response rates have been calculated annually for the ACS development program. Survey response rates are commonly used in assessing the potential for unit nonresponse error. These rates are traditionally calculated as the number of responding units divided by the total number of eligible units. Because the ACS design relies on three sequential data collection modes, the calculation of the survey response rate is slightly more complex.

The basic definition of the C2SS response rate follows the guidelines developed by the AAPOR and documented in the Statistical Policy Working Paper 31 guidelines.[11] Applying these guidelines to the C2SS design required some adaptations. As a result, two important points need to be highlighted. First, the data summarized for the C2SS come from all mail, CATI, and CAPI interviews received between January 1, 2000 and December 31, 2000, regardless of the month in

[11]American Association for Public Opinion Research, 2000. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys," Ann Arbor, MI: AAPOR.
"Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001.

sample. The data set includes noninterviews in this same time period. Second, for purposes of studying error introduced by unit nonresponse, the Census Bureau calculated weighted survey response rates. All interviews and noninterviews are weighted to reflect the C2SS sample, including the CAPI subsampling weights. The use of a weighted rate is consistent with the recommendation included in Statistical Policy Working paper 31.[12]

To aid in interpreting these weighted survey response rates, Table 2 provides distributions of interviews by mode. These three rates are calculated as the number of completed interviews collected by mail, CATI, and CAPI relative to the survey response rate. By referring to the percent of CAPI interviews, these rates allow data users to assess the proportion of the universe subsampled for CAPI. Unlike the operational response rates reported in Table 1 on page 10, which measured operational performance, these distributions of interviews across modes simply report the proportion of final interviews collected by mail, CATI, and CAPI.

Table 2: The C2SS Survey Response Data

	Rate (in percent)
Weighted Survey Response Rate	**95.4**
Mail interviews	56.2
CATI interviews	7.3
CAPI interviews	31.9
Noninterview rate	**4.6**

It is important to note that the survey response rate is insufficient as a measure of unit nonresponse error. At best they measure the potential for error (or bias) in the estimates produced from the C2SS attributable to unit nonresponse. The C2SS uses weighting techniques to account for noninterviews in an attempt to minimize the effect of unit nonresponse error. When the weighting assumptions regarding nonresponse characteristics are not fully met, nonresponse error will impact survey estimates. In the case of the C2SS, weighting uses data from responding households in the same tract and month as the nonrespondents. The degree to which nonrespondents may differ from respondents for the item of interest determines whether there is nonresponse error. Nonresponse error is magnified by the rate of nonresponse, which is why minimizing the nonresponse rate is crucial.

A survey response rate of 95.4 percent is considered an excellent rate of response. For example, the Current Population Survey (CPS) response rate, a monthly household survey, for the same timeframe was 93.1 percent.[13] Although the two rates are not readily comparable due to key definitional differences, the comparison is intended to provide a context for evaluating the C2SS rate. Like the decennial census long form survey, the C2SS is mandatory, leading to increased mail response rates and greater cooperation. Unlike the decennial census long form survey, the

[12] "Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001.
[13]The Current Population Survey (CPS) is a monthly sample survey of the U.S. population that provides employment and unemployment figures as well as current data about other social and economic characteristics of the population. The Census Bureau collects the data for the Bureau of Labor Statistics.

C2SS requires that all interviews be conducted with a household member. In the decennial census, if after repeated visits a household member cannot be reached, information about the nonresponding unit can be obtained from a neighbor, which is not allowed in the C2SS. This difference could result in higher levels of unit nonresponse in the C2SS versus the decennial census.

Evaluations are providing valuable information on nonresponse error

Special studies are shedding more light on the characteristics of the nonresponding universe in the C2SS. These studies provide some preliminary assessments of unit nonresponse error. Recent research addressed whether high survey response rates in the C2SS were maintained for all major racial and ethnic subgroups. Data from Census 2000 were used to sort all tracts in the nation into strata based on the percentage of the population in those tracts reporting a race or ethnicity of White, African American or Black, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, and Hispanic. The C2SS survey response rates were calculated for each stratum. Appendix II provides greater methodological details of this study.

Table 3 summarizes the C2SS survey response rates for the strata with 75 percent or more of the population reporting a specific race or ethnicity. The exceptionally small sample size for Native Hawaiian or Other Pacific Islander did not allow reliable estimates to be produced. Additional research is underway to produce more meaningful results for these groups. The data show some differences among strata, but the strata-level survey response rates remained consistently high. Tracts with high concentrations of African Americans or Blacks, and American Indians or Alaska Natives, have significantly lower survey response rates than those with high concentrations of Whites. This implies that the racial and ethnic characteristics of nonrespondents differ from the characteristics of respondents.

Table 3: The C2SS Survey Response Rates by Dominant Racial/Ethnic Group

Tracts with 75 percent or more of the population reporting a race or ethnicity of …	**Number of tracts**	**Weighted survey response rate**
White	34,315	95.9
African American or Black	3,642	92.4*
American Indian or Alaska Native	66	89.1*
Asian	93	95.2
Hispanic	1,646	95.8

*Significantly different from White at the 90 percent confidence level

Table 4 shows stratum-level rates of the percentage of interviews collected by mail, CATI and CAPI. As explained earlier, these rates reflect the ratio of the interviews by each mode to the survey response rate. The proportion of interviews collected by mail in predominantly African American or Black, American Indian or Alaska Native, and Hispanic areas were well below the rate for areas predominantly White—34.9, 16.6, and 34.2 percent versus 60.5 percent, respectively. Additional analysis is needed to determine if the low proportion of mail interviews might be due to problems in getting a questionnaire delivered in these areas. New methods to

reduce differential levels of mail response should be developed. The practical effect of low proportions of interviews being collected by mail and CATI is to increase the fraction of the original sample subjected to CAPI subsampling and thus to increase sampling error (discussed in a subsequent section). Based on these findings, the Census Bureau is developing options to oversample in areas of low mail response.

Table 4: The C2SS Interview Distributions by Dominant Racial/Ethnic Group

Tracts with 75 percent or more of the population reporting a race or ethnicity of …	Percent mail interviews	Percent CATI interviews	Percent CAPI interviews
White	60.5	7.3	28.1
African American or Black	34.9	8.9	48.6
American Indian or Alaska Native	16.6	2.6	69.9
Asian	58.6	4.1	32.5
Hispanic	34.2	8.3	53.3

*Significantly higher than White at the 90 percent confidence level

Additional evaluations are needed to understand the reasons for noninterviews. Gaining a better understanding of the types of noninterviews will allow the ACS to develop appropriate measures to improve response. A study is planned to analyze the detailed set of reasons that are reported for noninterviews, such as refusals, noncontacts, other noninterviews (due to such reasons as language barriers), and insufficient data.[14] A comparison to Census 2000 is planned to better understand the characteristics of nonrespondents in the C2SS. The C2SS CAPI noninterviews for the months of April, May, and June of 2000 will be compared to interviewed Census 2000 households. Analysts plan to summarize basic demographic characteristics including age, sex, race, Hispanic origin, relationship, and household size.

Special efforts are in place to minimize the occurrence of unit nonresponse

The C2SS included multiple programs designed to increase unit response. Additional efforts can target areas with the highest levels of unit nonresponse. In the mail mode, prenotice and reminder mailings were sent to all mailable sample addresses and a second questionnaire mailing package was sent to all addresses not responding to the initial questionnaire mailing. A telephone assistance line (in English and Spanish) provided assistance and encouragement for households to complete and return forms by mail. Experienced interviewers, who were often bilingual, conducted CATI and CAPI interviews.

[14] The C2SS received some mail returns that were determined after Edit Followup and final data processing not to have sufficient data to qualify as a completed interview. These cases were reclassified as noninterviews due to insufficient data.

Several improvements are underway to maintain and, where possible, increase these response rates. The Census Bureau has identified the need to improve rates of mail response for African American or Black, American Indian or Alaska Native, and Hispanic households. ACS designers plan to test a new set of letters and develop a program to provide questionnaires in Spanish. Development of a more user-friendly form may also improve mail response. Census Bureau staff will continue to work closely with advisory committees to develop targeted methods to improve CAPI response rates in areas experiencing the lowest rates, including those with high proportions of households speaking languages other than English and Spanish. The continuous monitoring of response rates provides an opportunity for the ACS to identify and address unit nonresponse issues.

Low levels of item nonresponse are critical to the success of the ACS

Item nonresponse occurs when a responding unit fails to provide complete and usable information for all required items. Item nonresponse will occur in all data collection modes, often for different reasons. A respondent may complete a mail form in error and omit sections or questions unintentionally. CATI and CAPI interviewers may find an otherwise cooperative respondent unwilling to provide them with personal information, such as income. The final ACS estimates that result from any item or battery of items can be adversely impacted when item nonresponse rates are high. To produce the most meaningful and accurate data, item nonresponse error in the ACS must be controlled. To accomplish this, the ACS must measure the rates of item nonresponse and determine possible explanations for why these data are not being collected. This line of documentation, analysis and research will enable the survey to target special efforts where they are needed to improve rates of item response.

The C2SS imputation rates for basic demographic items were significantly lower than those for Census 2000

The potential impact of residual item nonresponse in the C2SS, like most other surveys, is corrected through the use of imputation. For this reason, imputation rates are often used as a measure of the level of item nonresponse. Imputation rates are computed as the ratio of the number of eligible units or people that had a value imputed for that item to the number of units or people eligible to have responded to that item. This definition is consistent with the guidelines proposed in Statistical Policy Working Paper 31.[15]

The process by which values for missing or inconsistent data are determined is complex and varies by survey or census. In general, there is a continuum of certainty about the probable content of a missing data item. Analysts are confident about assigning values to a missing data item when related information is available on the same housing unit or person record. For example, first name may be used to assign a value of sex. Such imputations are often known as assignments since they do not rely on data from a separate record. Confidence is lower when values for missing or inconsistent items cannot be derived from the same housing unit or person

[15] "Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001.

record and must come from other persons in the household or other responding households or housing units believed to have similar characteristics. When such donors are used, the item is said to have been allocated. Such imputation is considered less accurate than using information derived from the housing unit or person record.

The decennial census and the C2SS use similar data processing methods to deal with item nonresponse. Tallies of the edit results from the C2SS and Census 2000 were the basis for the calculations of item imputation rates. For population items, the imputation rate is defined as the total number of allocations for an item as a percent of the total population eligible to report that item. Housing item imputation rates and potential reasons for any differences observed are not discussed but will be included in the next report describing basic demographic item comparisons.

Table 5 summarizes item imputation rates for the population items included in both the C2SS and Census 2000. A brief description of these items can be found in Appendix III. Because the ACS is intended to replace the decennial long form, comparisons to the long form data items will be more meaningful and are the subject of future reports.

Unit nonresponse is not reflected in these rates. Only respondents are included in the denominator for these rates. The census population item imputation rates are based on the Census 2000 household population (excludes the group quarters population) that met the minimum edit requirement for population data.[16] The census results shown are all population items from both the short and long forms. Unlike the C2SS, there was no content followup for the Census 2000 long form survey. The C2SS item imputation rates are based on the C2SS estimates of household population before the noninterview adjustments and population controls have been applied. These results show lower imputation rates for the C2SS on all population items.

As with weighting methods, imputation techniques do not completely compensate for these missing data. Some unknown level of error remains in the data due to item nonresponse. Imputation methods rely on models for generating values for missing items. Documentation of C2SS imputation methods will be provided in future reports.

Table 5: Comparison of Population Item Imputation Rates[17]

	Census 2000 (percent of eligible items)	**C2SS (percent of eligible items)**
Relationship	2.2	1.5*
Gender	1.0	0.5*
Age	3.6	2.4*
Hispanic Origin	4.2	3.6*
Race	3.9	2.4*

*Significantly different from Census 2000 at the 90 percent confidence level

[16] The minimum edit requirement for population data requires the presence of at least one data-defined person. That is, there are at least two basic population items defined.

[17] This computation excludes item nonresponse that was addressed through assignment methods.

Research and evaluation will provide insight into reducing levels of item nonresponse

Research and analysis on a number of fronts can contribute to a greater understanding of the factors associated with item nonresponse. Such analyses include examining current and alternative imputation methods and conducting empirical analysis, leading to operational improvements.

The C2SS Edit Followup operation provides empirical data on items that respondents fail to complete correctly on their mail returned forms. Analysis of these data will identify problem questions that might benefit from redesign. For example, questionnaire design changes (format, presentation, wording) may be warranted to reduce levels of item nonresponse. Additionally, current analysis of imputation rates for all items by mode of data collection allows for the recognition of those questions that are experiencing high levels of CATI and CAPI item nonresponse. Such information can be used to identify questions that may require additional interviewer training, changes in questionnaire processing, or changes in the actual wording of the questionnaire. Improvements in the effectiveness of the Edit Followup operation should also reduce final levels of item nonresponse.

Additionally, research is currently taking place to evaluate item nonresponse for households with limited English-speaking proficiency. This research is providing valuable information to support the development of language assistance programs.

Methods designed to maintain high levels of item response could be improved

In the C2SS, special followup efforts were used to reduce levels of item nonresponse. Questionnaires returned by mail were reviewed for completeness and consistency. Questionnaires that failed this review were sent to a telephone followup operation to recontact the household and obtain the missing information. The first released report on ACS operational feasibility identified that this followup operation had production problems.[18] Therefore, the full intent of this operation was not realized. Item nonresponse rates for mail returned questionnaires in the C2SS are likely to be higher than they would have been if production problems had not occurred. Operational analysis is needed to improve the effectiveness of this edit and followup operation to ensure high quality data for mail returned questionnaires.

CATI and CAPI instruments use built-in edits to identify inconsistent responses. Experienced interviewers have developed techniques to encourage respondents to provide more complete responses. A greater understanding is needed of how CATI and CAPI interviewers probe to obtain complete item responses. CATI and CAPI interviewers often experience problems in convincing respondents to provide data for sensitive items, such as income. Specialized training may increase the response rates for such items.

[18] "Meeting 21st Century Demographic Data Needs—Implementing the American Community Survey: Demonstrating Operational Feasibility," U.S. Census Bureau, July 2001.

Sample Completeness Ensured Accurate Survey Representation

Survey undercoverage and unit nonresponse lead to the same problem—excluding a certain group of people or households from the survey.[19] In the C2SS, low levels of survey undercoverage and high survey response rates are needed to ensure that survey estimates accurately represent all population groups. The practical effect of both survey undercoverage and unit nonresponse is that C2SS estimates may be biased if data are based on incomplete information. For example, if African American or Black males between the ages of 18 and 24 with incomes below $20,000 are disproportionately not in the sample, the C2SS could report incorrectly on the characteristics of African American or Black males (e.g., too high an income for the African American or Black population, as well as for the total population).

Estimates of C2SS sample completeness are consistent with the sample completeness measures from the 1990 census long form survey

Sample completeness measures the extent to which the interviewed survey sample represents the decennial census population. Differences can occur due to survey undercoverage, unit nonresponse, and differences in the address frame used for sample selection. When the people included in the C2SS or the long form sample are different from those in the census count, estimates of characteristics can be affected, even though the sample estimates of total population are controlled to the decennial census benchmark. From a practical standpoint, the impact of completeness in the address frame may not be fully reflected in the sample completeness ratio, since the frame used in the survey and the decennial census is essentially the same—the MAF. Therefore, the completeness of the MAF affects both the survey and the census benchmark.

To evaluate the C2SS, sample completeness was evaluated in relation to Census 2000. Specifically, C2SS weighted population estimates—without adjustments for nonresponse or coverage error—were divided by the population counts from Census 2000. A similar ratio was calculated for the 1990 long form relative to the 1990 decennial census counts. Sample completeness measures for the Census 2000 long form will be produced when the data are available. Values closest to 1 indicate the lowest levels of difference, relative to the census. Appendix IV provides details of the methodology used to produce these results.

As seen in Table 6, the C2SS sample completeness ratios for the total population were comparable to the ratios for the 1990 decennial long form. Significant differences between the two estimates are denoted with an asterisk. A ratio of 0.902 indicates that the C2SS represented

[19] Survey undercoverage in the C2SS cannot be compared to decennial census undercoverage. Census undercoverage is evaluated using an independent post-enumeration survey, such as the Census 2000 Accuracy and Coverage Evaluation (ACE) survey. The ACE results were compared to the census results to identify whole household and within household differences, indicating census coverage errors. Survey undercoverage is evaluated by comparing aggregate estimates from a survey to census counts or intercensal estimates derived from census counts. Therefore, measures of survey undercoverage will not reflect any housing units or persons that are missed in both the census and the survey.

about 90 percent of the 2000 census population—10 percent were not represented. Similarly, the 1990 census long form represented about 90 percent of the 1990 census population.

Table 6: Comparison of Sample Completeness Ratios

	C2SS to Census 2000	**1990 Long Form to 1990 Census**
Total Population	0.902	0.897*
Hispanic (any race)	0.881	0.845*
Not Hispanic	0.905	0.903

*Significantly different from the 1990 Census at the 90 percent confidence level

As mentioned earlier, there are numerous explanations for this 10 percent difference. C2SS nonresponse, which occurs when efforts to collect data by mail, CATI, and CAPI are unsuccessful, contributes to this shortfall. For these sample cases, no data are collected. In addition, the levels of sample completeness in the C2SS can result from differences in the content of the MAF used in the C2SS versus in Census 2000. The C2SS and Census 2000, in fact, used different versions of the MAF due to timing constraints. Since the sample completeness measures are relative to the decennial census, errors such as under and over coverage in the decennial census, affect both the estimated ratios and their interpretation.

The impact of sample completeness on the representativeness of the sample depends on whether the people who are included in the census, but not in the C2SS, have different characteristics. This is difficult to determine, since survey data for these people are, by definition, not available. The best assurance of having a representative sample is to have a completeness ratio close to 1.

Table 7 provides additional sample completeness measures by race. For most groups, the C2SS ratios compare favorably to the 1990 decennial long form ratios. However, the interpretation of these ratios must acknowledge possible differences in the collection of race responses. Some portion of the difference between the C2SS and Census 2000 is likely due to nonresponse and to survey undercoverage. But some of the differences are due to race reporting differences between the decennial census and the C2SS. The concept of race is complex and it appears that very minor differences in how the data are collected can affect the responses to a far greater degree than previously understood. This situation has resulted in plans for extensive testing of race questions for future use in the ACS and the 2010 census.

Table 7: Additional Comparisons of Sample Completeness Ratios

	C2SS to 2000 Census	1990 Long Form to 1990 Census
White	0.935	0.912*
Black	0.851	0.807*
American Indian/Alaska Native	0.881	0.914
Asian	0.892	0.894
Native Hawaiian / Other Pacific Islander	0.882	0.872
Some Other Race	0.607	0.863*

*Significantly different from the 1990 Census at the 90 percent confidence level

MAF modernization and other activities could improve sample completeness

Several important activities are intended to improve the sample completeness of the ACS. An evaluation study comparing C2SS and Census 2000 records is underway and will provide important information on survey undercoverage and nonresponse in the C2SS. Additionally, research on providing greater language assistance may lead to improved within household coverage. Finally, the MAF/TIGER modernization program is a multi-million dollar effort intended to not only modernize the hardware and software infrastructure but to also ensure a continuously updated and reliable address and geographical system for the Census Bureau's censuses and surveys. The effort includes a MAF/TIGER evaluation program for providing improved measures of housing unit coverage and other accuracy indicators. Without a successful modernization program ensuring a complete, accurate MAF over the decade, ACS coverage is likely to suffer.

Processes Exist to Control Measurement and Processing Errors

Measurement and processing errors are two additional components of accuracy, affecting the quality of the C2SS. Measurement error refers to the difference between the observed value of a variable gathered during data collection and the true, unobserved value of the variable. Measurement error includes response error, which occurs if a respondent does not understand the meaning of a question or fails to recall the information accurately. Interviewer error can also be a source of systematic measurement error if interviewers are not properly trained or if they misinterpret their procedures.

Processing error occurs during the series of operations that convert reported data to consistent machine-readable information and published estimates. For example, error will be introduced if a data entry clerk keys the wrong information during data capture or a CATI or CAPI transmission error occurs. In addition, clerical coding is needed for some items and coding errors are possible. In the C2SS, a detailed set of edits is used. Errors introduced through these edits are another possible source of processing error.

Accurately assessing the extent and nature of measurement and processing errors and determining how to minimize them is a difficult task requiring an ongoing research and testing program. ACS developers are learning from the C2SS experience to better understand potential errors and to strengthen procedures for reducing them.

ACS methods currently include procedures to control errors

As reported in the first ACS report on operational feasibility, some procedures have been put into place to help control measurement and processing errors in the C2SS.[20] These procedures are either inherent to the use of ACS methods or are applied as part of the C2SS quality assurance activities. Such procedures have the potential to result in lower measurement and processing error levels than are found in the decennial census long form survey.

The CATI and CAPI operations benefited from several quality assurance activities. Because both CATI and CAPI use a computer to conduct interviews, the software has numerous checks and edits in place to ensure accuracy. For example, the software prevents most errors such as out-of-range responses or skipped questions. Additionally, in CATI, monitoring is used to check for other interviewer errors such as asking the questions incorrectly or keying a different answer from what the respondent provided. A formal quality control reinterview program is built into the CAPI operation. The work of field interviewers is sampled and respondents are recontacted to determine if there is any evidence of falsification or other substandard performance.

To help ensure that processing errors are not introduced during keying, a detailed quality assurance program has been developed. [21] The quality assurance process is designed to keep work unit total error rate below 1.5 percent, which prevents keying from being a serious source of error.

[20] Ibid.
[21] Ibid.

Subject matter experts review edited data to minimize errors being introduced by the edit and imputation procedures.

As seen from these examples, ACS methods and quality assurance are helping to control measurement and processing errors in the C2SS. However, it is critical that error continue to be measured and, when appropriate, procedures changed to reduce measurement and processing errors.

Efforts are underway to identify potential sources of measurement and processing errors

Benefiting from the C2SS nationwide implementation, ACS analysts have been implementing several research and testing activities that are designed to better understand and reduce measurement error. To determine if additional training is needed in the collection of race data, an interviewer-debriefing project is underway. Focus groups are being used to assess how improvements could be made to the current CAPI instrument to improve the quality of data collected during interviews in languages other than English. Additional work is needed to more clearly understand the effects of measurement and processing error. This is especially important given the substantial increase in sample size beginning with full ACS implementation in 2003. Such an increase is likely to *initially* increase errors in the survey until new interviewers and staff become more experienced. Consequently, additional monitoring, as well as research and analysis, are called for to help offset any potential increases in errors.

Sampling Error was Controlled and Will Be Slightly Higher Than the Decennial Census Long Form Survey

In addition to the nonsampling errors discussed above, another critical measure of survey quality is the level of sampling error. As discussed earlier, beginning in 2003, the ACS will sample 2.5 percent of the population annually, spread across all geographic areas. Each of the ACS test sites has such geographically spread samples, with sampling rates of 3 or 5 percent in most sites and 1 percent in the Houston site. The C2SS sample has a different design, in which each state's sample is spread only over a sample of counties—not spread over all geographic areas. Therefore, analysts used the data from the ACS test sites, rather than from the C2SS, to examine sampling errors. While the targeted levels of sampling error for single-year estimates were met overall, differentials in levels of mail response for some population groups indicate that sampling error is disproportionately higher, suggesting the need for design changes. Further, it is important to note that, due to smaller sample sizes, all ACS estimates will have slightly higher levels of sampling error than the decennial census long form survey.

Sampling error provides a measure of survey precision

Sampling error refers to the variability that occurs by chance because a sample—rather than all units in a population—is surveyed. In general, the larger the sample, the smaller the sampling error. Anything, in effect, that reduces the sample size, increases sampling error.

Sampling error can generally be computed in a straightforward manner. However, as cost, schedule, or performance constraints increase, the sample design can become increasingly complex. This complexity translates into additional methods needed to compute sample estimates and associated measures of sampling error. A measure of sampling error is the variance or standard error of a statistic. The impact of the standard error depends on its magnitude relative to the size of the estimate. For example, a standard error of 10 is relatively small compared to an estimate of 5,000 but quite large for an estimate of 50. To get at this relationship, a related, but different statistic—the Coefficient of Variation or CV—is computed. The CV is defined as the ratio of the standard error of the estimate to the value of the estimate. Therefore, the smaller the CV, the more precise the estimate will be. For example, one expects that for a random sample of respondents, roughly 50 percent will be male. If the standard error is 1.5 and that number is divided by 50, the CV is 3 percent of the estimated value. Said another way, the relative measure of sampling error compared to the estimate for the item "male" is 3 percent.

Coefficients of Variation for single-year estimates were generally consistent with targeted level

Based on data from the 1999 ACS test sites, CVs were computed indicating sampling error levels met survey objectives—a trade-off between cost and accuracy. The most widely discussed numerical objective was for single-year estimates for a population of 65,000. This objective was chosen as the limit at which data users will use multiple instead of single-year data. Specifically,

the expectation was that estimates for a population of 65,000 for items occurring in 10 percent of the population will yield CVs of about 12 percent. The nomenclature "10 percent data item" refers to the occurrence of a data item, say "poverty" in the population. These 10 percent data items, known as regular items, include age, school enrollment, marital status, and employment. Depending on the occurrence in the population, different data items generally will have higher or lower anticipated CVs. Two types of items were expected to have higher CVs in the ACS than regular items—correlated and vacant data items.

Correlated items have responses that are usually concentrated within the same household and thus have higher CVs than other items in the C2SS. These include race, Hispanic Origin, ancestry, language spoken at home, and citizenship status. These items have higher CVs because within household clustering reduces the number of independent observations that the survey is making about a particular characteristic. One household member's response to, say, "race" often means other members will have the same response. Correlated items also have been observed to have higher CVs than regular items for the decennial census long form in past censuses.

The higher CVs for vacancy-related items are specific to the C2SS because of the subsampling for CAPI nonresponse followup. Estimates of vacant housing unit characteristics tend to be interviewed by CAPI rather than by mail or CATI. Consequently, vacant housing characteristics are almost always subjected to the one-in-three CAPI subsampling, increasing the CVs.

Table 8 presents mean and median CVs for these three types of data items. See Appendix V for more details on the methodology used to produce them.

Table 8: Summary of CVs by Type of Data Item

Type of Data Item	Mean CV (in percent)	Median CV (in percent)
Regular	12.2	11.8
Correlated	19.3	17.6
Vacant	16.5	14.0
Total	14.3	12.5

Based on the 1999 ACS test sites, the actual CVs for regular data items (12.2 percent mean CV) are consistent with target CVs of 12.0 percent. As anticipated, the CVs for the correlated and vacant items were higher than those for the regular items. Further examination is needed to determine if these target CVs are maintained for all population groups.

Differential levels of mail response imply higher CVs for some population groups

Actual CVs for the full range of characteristics for all population groups have not yet been analyzed. The ACS sample design includes a subsample of those who do not respond by mail or telephone and must be interviewed in person. This subsampling of one-in-three of all respondents eligible to be interviewed in person was selected as an optimum trade-off between cost and

sampling error. However, applying it randomly to all nonresponding households means that a larger proportion of some population groups with a lower-than-average mail response rate become eligible for CAPI subsampling. Hence, such groups will tend to have a smaller sample size and, therefore, greater sampling error.

Table 4 on page 17 (The C2SS Interview Distributions by Dominant Racial/Ethnic Group) illustrates that when compared to census tracts with high White populations, a lower proportion of interviews are collected by mail in census tracts with high African American or Black, American Indian or Alaska Native, and Hispanic populations. Based on these results, sample will be reallocated from tracts with higher proportions of interviews collected by mail to allow more nonresponse followup sample in below-average tracts. Alternative design options to assure more uniform CVs are being examined. Even after a more uniform level of sampling error is achieved among population groups, because of differences in sample size, sampling error will be higher for the ACS than for the decennial long form survey.

CVs for small areas in the ACS are expected to be slightly higher than the decennial census long form survey

CVs differ based on the size of the estimate and the size of the sample, changing according to standard statistical formulas. So a survey that gives a 12 percent CV for an item that is*10 percent of the population* typically will have a smaller CV of about 4 percent for a item that is *50 percent of the population*. Similarly, if the 12 percent CV for a 10 percent item is for an estimate based on *1 year of sample*, the corresponding estimate based on *5 years of sample* will have a smaller CV of about 5 percent. The target criterion—a 12 percent CV for an estimate that is 10 percent of the population—therefore implies target CVs for a range of other estimates and for 1-year, 3-year, and 5-year samples.

Due to the somewhat smaller sample sizes, the ACS 5-year averages are expected to have somewhat higher CVs than corresponding Census 2000 long form estimates. The ACS 5-year sample starts out at 15 million addresses, compared to about 19 million for the Census 2000 long form, and is then reduced further by subsampling for nonreponse followup. On the other hand, the long form may lose more sample units to nonresponse than the ACS. Census Bureau analysts have predicted the net effect to be that CVs for the ACS for 5-year averages will have 1/3 higher CVs than the CVs for the decennial census long form for the same estimate. In other words, the ACS CVs will be 1.33 times as large as the CVs for Census 2000. This prediction will be studied further when the Census 2000 long form CVs are available.

The premise of the ACS design is that this moderate increase in CVs for a 5-year average is worthwhile in order to obtain regular updates of the estimates throughout the decade, and to obtain what is expected to be a generally lower level of nonsampling error. Since the variations in the mail response rates can cause the ACS CVs to be higher than this target in areas with low mail response, it is important to take steps to maintain the target CVs uniformly in all areas.

Conclusions

Assuming congressional approval and funding, the ACS will be implemented beginning nationwide in 2003, replacing the once-a-decade decennial census long form survey. While the decennial long form survey has been the standard for detailed demographic and socioeconomic data, the ACS should improve the overall quality of such data. Because it is ongoing, the ACS can provide more timely data and opportunities to improve methods and reduce error. While sampling error will be slightly increased, nonsampling error should be decreased when compared to the decennial census long form survey.

This report documents the C2SS performance examining timeliness but primarily focusing on the quality dimension of accuracy. Additional key measures of accuracy for detailed demographic and socioeconomic data items will soon be evaluated. Overall accuracy was high for the C2SS, which is an early indication that the ACS will have the quality to replace the long form. Maintaining the C2SS quality levels, however requires (1) funding to continue implementing C2SS methods, (2) dealing with *initial* increases in nonsampling error that could result from workload and staffing increases when first implemented nationwide, and (3) resources directed to monitoring and evaluating data quality.

Some corrective actions are currently underway. For example, ACS analysts are examining the sample design to determine which different sample allocation should be employed to obtain more uniform sampling errors across population groups. Additionally, the MAF/TIGER modernization effort is underway; and, if successful, the outcome should improve survey coverage. The continuous nature of the survey coupled with a systematic assessment and evaluation plan provides multiple opportunities for refining and improving methods to address deficiencies and adapt to change over the decade as the nation continues to change. However, to systematically identify such changes and the opportunities they present, adequate resources must be allocated to fully examine the wide array of complex and interdependent error components comprising the ACS.

While important, the results provided in this report do not yet give a complete picture of the performance of the C2SS. Comprehensively depicting the dimensions of quality—relevance, accessibility, timeliness, and accuracy—will be the subject of future assessments and reports.

Appendices

Appendix I: Definitions of the C2SS Survey Response Rates, Interview Distributions, and Operational Response Rates

1. Background

The American Association for Public Opinion Research suggests that survey cases can be categorized into the following four groups: interviewed cases, eligible cases that are noninterviews, cases of unknown eligibility, and cases that are not eligible to be interviewed.[22] Response rates are calculated as:

$$\frac{I}{I + R + NC + O + eU}$$

Where
I = the number of interviewed cases
R = the number of eligible cases that are noninterviews due to refusals
NC = the number of eligible cases that are noninterviews due to noncontacts
O = the number of eligible cases that are noninterviews due to other reasons for noninterviews
U = the number of cases of unknown eligibility that are not interviewed
e = the estimated proportion of sample cases of unknown eligibility that are eligible

2. Survey Response Rates and Interview Distributions

This logic guided the definitions of the C2SS survey response rates as well as the operational response rates for each of the three data collection modes.

Let,
I(m) = the number of cases completed by mail
I(t) = the number of cases interviewed by CATI
I(p) = the number of cases interviewed by CAPI
R = the number of eligible cases that are noninterviews due to refusals
NC= the number of eligible cases that are noninterviews due to noncontacts (e.g., unit was never contacted prior to closeout)
O = the number of eligible cases that are noninterviews due to other reasons for noninterviews (e.g., language problems, unresolved callbacks, insufficient data)
U = the number of cases of unknown eligibility that are not interviewed (e.g., cases that the interviewer is unable to locate, cases with missing status codes).
e **=** the estimated proportion of sample cases of unknown eligibility that are eligible (i.e., e was assigned a value of 1).

[22] Association for Public Opinion Research. 2000. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys," Ann Arbor, MI: AAPOR

The ACS and the C2SS select a sample of addresses each month. Data collection for each monthly panel takes place over a three-month period. Mail returns are collected in the first month, CATI interviews and additional mail returns in the second month, and CAPI interviews and some late mail returns in the third month. To facilitate accurate response, data are collected as of the time of interview rather than as of the time of initial mailout. For example, the cases interviewed in January consist of the early mail returns from the January sample panel, the CATI interviews and additional mail returns from the previous December sample panel, and the late mail returns and CAPI interviews from the previous November sample panel. For purposes of survey estimation, the final January sample consists of these cases whose data are collected in January, and any noninterviews from the CAPI workload in January (November sample panel). Mail returns received in January and later determined to have insufficient data are classified as noninterviews for January.

Initial weights were applied to the C2SS sample units to account for differences in selection probabilities that resulted from the design of the sample. This includes CAPI subsampling weights.

$$\text{Survey response rate} = \frac{I(m) + I(t) + I(p)}{I(m) + I(t) + I(p) + R + NC + O + U} * 100$$

$$\text{Percent mail interviews} = \frac{I(m)}{I(m) + I(t) + I(p) + R + NC + O + U} * 100$$

$$\text{Percent CATI interviews} = \frac{I(t)}{I(m) + I(t) + I(p) + R + NC + O + U} * 100$$

$$\text{Percent CAPI interviews} = \frac{I(p)}{I(m) + I(t) + I(p) + R + NC + O + U} * 100$$

3. Operational Response Rates

The following rates were calculated to measure operational performance. In contrast to the survey response rates, they are based strictly on operational data for the sample months of January through December of 2000.

$$\text{Mail response rate} = \frac{I(m)}{I(m) + R(m) + NC(m) + O(m) + eU(m)}$$

RC(m) + NC(m) + O(m) = the total number of eligible cases that were mail noninterviews. We can not distinguish between refusals, noncontacts and other noninterviews. Sample cases that were never mailed out were considered not eligible. No cases were considered to have unknown eligibility.

$$\text{CATI response rate} = \frac{I(t)}{I(t) + R(t) + NC(t) + O(t) + eU(t)}$$

R(t), NC(t), O(t) refer only to the noninterviews recognized during CATI. Sample cases with no available phone number or a confirmed bad phone number are considered not eligible for CATI interviewing. Cases with unknown eligibility (e.g., ring-no-answer) were counted as eligible noninterviews.

$$\text{CAPI response rate} = \frac{I(p)}{I(p) + R(p) + NC(p) + O(p) + eU(p)}$$

R(p), NC(p), O(p) refer only to the noninterviews recognized during CAPI. Sample cases determined to be commercial, nonexistent or demolished are considered not eligible for CAPI interviewing. Cases with unknown eligibility (e.g., cases that the interviewer could not locate) were counted as eligible noninterviews.

Appendix II: Methodology for Review of C2SS Response Rates by Race and Ethnicity

1. Methodology

a. Sample Design

A total of 14 monthly panels contribute to the C2SS final annual sample. The data summarized for the C2SS come from all mail, CATI, and CAPI interviews as well as all noninterviews recognized between January 1, 2000 and December 31, 2000. Over 800,000 housing units in over 53,000 census tracts made up the C2SS universe.

Census 2000 data were used to identify all tracts where the percentage of persons who reported a race of White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, and Hispanic was greater than 75. For purposes of this analysis the race could have been reported alone or in combination with another race. Table 1 summarizes, for each racial and ethnic group, the number of tracts that met this requirement. Table 1 also provides the final housing unit sample found in the combined set of tracts. The sample size for the Native Hawaiian or Other Pacific Islander group is too small to produce meaningful results. Although small, the sample sizes for the American Indian or Alaska Native and the Asian groups will allow for a preliminary review of the rates.

Table 1: Summary of Sample

Racial or Ethnic Group	Number of Tracts	Number of Housing Units
White	34,315	584,068
Black or African American	3,642	36,011
American Indian or Alaska Native	66	1,640
Asian	93	1,717
Native Hawaiian or Other Pacific Islander	3	39
Hispanic	1,646	18,673

b. Calculation of Survey Response Rates

Weighted survey response rates were calculated for each of the six racial and ethnic groups. The survey response rate is defined as the ratio of all completed interviews (across all modes) to the combination of interviews and noninterviews. This definition is consistent with the guidelines developed by the American Association for Public Opinion Research. [23]

I(m) = the number of cases interviewed by mail
I(t) = the number of cases interviewed by CATI
I(p) = the number of cases interviewed by CAPI

[23] American Association for Public Opinion Research. 2000. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys," Ann Arbor, MI: AAPOR

R = the number of eligible cases that are noninterviews due to refusals [24]
NC = the number of eligible cases that are noninterviews due to noncontacts[25]
O = the number of eligible cases that are noninterviews due to other reasons for noninterviews[26]
U = the number of cases of unknown eligibility that are not interviewed
e = the estimated proportion of sample cases of unknown eligibility that are eligible (i.e., ***e*** was assigned a value of 1).

$$\text{Response rate} = \frac{I(m) + I(t) + I(p)}{I(m) + I(t) + I(p) + R + NC + O + eU}$$

Sample cases determined to be commercial, nonexistent, or demolished are considered not eligible for survey interviewing. Cases with unknown eligibility (e.g. unable to locate) were all counted as eligible. Appropriate weights were applied to account for differences in selection probabilities. This includes CAPI subsampling weights. Standard errors were calculated along with 90 percent confidence intervals.

b. Calculation of Distributions by Mode of Data Collection

Distributions of interviews across modes were calculated for each of the six racial and ethnic groups. They are defined as follows:

$$\text{Percent mail interviews} = \frac{I(m) * 100}{I(m) + I(t) + I(p) + R + NC + O}$$

$$\text{Percent telephone interviews} = \frac{I(t) * 100}{I(m) + I(t) + I(p) + R + NC + O}$$

$$\text{Percent personal visit interviews} = \frac{I(p) * 100}{I(m) + I(t) + I(p) + R + NC + O}$$

These rates were also weighted to account for differential selection probabilities, and 90 percent confidence intervals were calculated.

[24] These are final CAPI refusals only. Mail and CATI refusals are eligible for CAPI.
[25] These are final CAPI noncontacts. Mail and telephone noncontacts are eligible for CAPI.
[26] Includes mail, telephone or personal visit interviews determined during final processing to have insufficient data to qualify as a complete or sufficient partial interview. Also includes CAPI noninterviews due to language barriers and specific other reasons.

Appendix III: Description of Population Items[27]

The following five items are described in this appendix: Relationship, Gender, Age, Hispanic Origin, and Race.

ITEM	QUESTION	RESPONSE CHOICES
Relationship	How is this person related to Person 1**?**	Husband or wife Son or daughter Brother or sister Father or mother Grandchild In-law Other relative Roomer, boarder Housemate, roommate Unmarried partner Foster child Other nonrelative
Gender	What is this person's sex?	Male Female
Age	What is this person's date of birth and what is this person's age?	Month, Day, Year of Birth Age (in years)
Hispanic Origin	Is this person Spanish/Hispanic/Latino? Mark (X) the "**No**" box if **not** Spanish/Hispanic/Latino	**No**, not Spanish/Hispanic/Latino Yes, Mexican, Mexican Am., Chicano Yes, Puerto Rican Yes, Cuban Yes, other Spanish/Hispanic/Latino – *Print group*
Race	What is this person's race? Mark (X) one or more races to indicate what the person considers himself/herself to be.	White Black, African Am., or Negro American Indian or Alaska Native – *Print name of enrolled or principal tribe.* Asian Indian Chinese Filipino Japanese Korean Vietnamese Other Asian – *Print race* Native Hawaiian Guamanian or Chamorro Samoan Other Pacific Islander – P*rint race below* Some other race – P*rint race below*

[27] The American Community Survey, Form ACS-1(2000)

Appendix IV: Calculations of the Sample Completeness Ratio

For the C2SS ratios:

$$\text{SC Ratio} = \frac{\sum W_{C2SS}}{C}$$

Where,

SCRatio = the sample completeness rate for the C2SS
$\mathbf{W_{C2SS}}$ = the initial sampling weight based on the C2SS sampling strata multiplied by the CAPI subsampling factor
C = the Census 2000 household population count.

The initial sampling rate for the C2SS takes into account the higher sampling rate in small states, the generally higher sampling rates that apply to the ACS comparison counties that are used in the C2SS estimates, and the especially high rates for small governmental units in the ACS sites.

To produce the overall sample completeness ratio, the population estimate for the C2SS prior to noninterview and population control adjustments is computed and divided by the Census 2000 household population count. The population estimate is computed by summing the weights for all of the people in C2SS interviewed households. Likewise, to produce sample completeness ratios for a specific population group, the population group estimates for the C2SS prior to noninterview and population control adjustments are divided by the Census 2000 household population count for that population group. The population estimates are computed as the sum of the weights for all people in a specific population group who were interviewed in the C2SS. If there were no nonresponse in the C2SS and if the C2SS had no coverage errors that were not also present in the census population counts, then the ratio would equal 1.

Standard errors for the sample completeness ratios were calculated directly for the C2SS ratios based on replicates. The actual method used is known as the standard successive differences replication methodology.

For the 1990 long form ratios:

$$\text{SC Ratio} = \frac{\sum W_{LF}}{C}$$

Where,

SCRatio = the sample completeness ratio for the 1990 long form
$\mathbf{W_{LF}}$ = the initial sampling weight based on the sampling rate used in the 1990 Census (1 in 2, 1 in 6, or 1 in 8)
C = the 1990 census household population count.

To produce the overall 1990 Census long form sample completeness ratio, the population estimate from the long form survey prior to noninterview and population control adjustments is computed and divided by the 1990 Census household population count. The population estimate is computed by summing the weights for all people interviewed in the Census on a long form. Long form households for whom responses were imputed to make up for a shortfall of actual long form cases were removed from the calculations. Population group estimates were produced comparably to the C2SS. Standard errors for the 1990 long form ratios were approximated based on generalized variance factors.

For the Race ratios:

Census 2000 and the C2SS allowed respondents to indicate more than one racial category for the first time. This was not true in 1990. Because of the multiple race responses, it was not possible to exactly compare the ratios for racial and ethnic groups in the C2SS and the 1990 Census. For purposes of this analysis, multiple race respondents in the C2SS and Census 2000 were randomly assigned to one of their indicated race groups (with probability 1/2 for each of 2 races, 1/3 for each of 3 races, etc.). All calculations were made based on this recoding of the race variable. In addition, in Census 2000 and the C2SS, Asians and Native Hawaiian and Other Pacific Islander were separate race categories. In the 1990 Census, they were combined. Detailed data from 1990 were used to distribute persons into the separate Asian and Native Hawaiian and Other Pacific Islander race categories.

Appendix V: Methodology for Computation of Coefficients of Variation (CV)

The 1999 American Community Survey (ACS) test sites included 36 counties. The 36 counties were grouped into 31 sites for publication, with single-year estimates being published for 21 of the 31 sites. The counties were selected so that there was at least one county in each of 24 groups or strata representing a combination of county population, difficulty of enumeration, and population growth between 1990 and 1995. The choice of comparison sites also attempted to balance areas by region, racial and ethnic representation, highly seasonal populations, migrant workers, American Indian reservations, and predominant occupation or industry type.

The 1999 ACS sample is a systematic sample of addresses from the August 1998 Master Address File (MAF) for the survey sites. The target sampling rate varied by block and depended on the site's population and characteristics of the block—whether the block was in a small governmental unit or in a large census tract.

Tabulation areas vary significantly in size, percentages for the item, and by sample size. The 1999 ACS test site design used three different sampling rates (1 percent, 3 percent, and 5 percent). So that as single table can show how CV's for different items and population sizes compared to the requirements, each standard error was converted to its underlying design factor. Then the design factor was used to derive what the CV would have been for a 2.5 percent sample for an estimate for a population of 65, 000 for items occurring in 10 percent of the population. Table 8 on page 27 provides mean and median CVs for all tabulation areas in the 21 sites for which data were published.

The design factor is a standard device used in sampling theory to compare the standard errors for estimates of different sizes based on different sampling rates. The computation was based on the 1999 data from the 21 sites published in 2000. The design factor can be calculated from the following formula:

$$\text{Standard error of the estimate} = \text{design factor} * \sqrt{\frac{97.5}{2.5} * 6500 * \left(1 - \frac{6500}{65{,}000}\right)}$$

The CV is estimated then as follows: $CV = \dfrac{\text{Standard Error of the Estimate}}{\text{Estimate}}$

Appendix VI: Glossary of Terms

The following represents a selection of terms that are frequently used in this report. Sources for this glossary are footnoted as appropriate and include documents internal to the Census Bureau as well as an external document published by the Congressional Quarterly, Inc. *Where ever possible, definitions are consistent with the Census Bureau Standard: Definitions for Survey and Census Metadata. Items not footnoted are defined in this report.* While not intended to be comprehensive, this glossary provides the reader with many of the concepts and phrases used throughout this report.

2000 Supplementary Survey (C2SS). [28] A Census 2000 experiment demonstrating that the American Community Survey (ACS) can be implemented nationwide at the same time, but separate from, the decennial census. Conducted in 1,203 counties, it is the test vehicle for reporting on the operational and technical performance of the ACS.

2001 Supplementary Survey. The second demonstration nationwide test. It is designed to report on the *usability and reliability* of collecting long form data using the ACS questionnaire and methods. Its data collection activities are continuing in the same 1,203 counties as its predecessor, the C2SS.

Accessibility.[29] The availability of information from the holdings of the agency, also taking into account the suitability of the form in which the information is available, the media of dissemination, the availability of metadata, and whether the user has a reasonable opportunity to know the data are available and how to access it.

Accuracy.[30] The difference between the sample result and the true population value. Attributes are measured in terms of sources of error (e.g., coverage, sampling, nonresponse, response, measurement, and processing).

American Community Survey (ACS).[31] The replacement for the decennial long form survey. When fully implemented in FY2003, it will collect the detailed demographic data traditionally collected on the decennial census long form from 3 million households a year, located in every county, American Indian and Native Alaskan area, and Hawaiian Homeland, as well as Puerto Rico. These data will provide updates on detailed characteristics about our nation every year, rather than only once every ten years. Implementation of the ACS will enable the 2010 census to collect only short form information.

[28] Census Bureau, "Meeting 21st Century Demographic Data Needs-Implementing the American Community Survey: July 2001, Report 1: Demonstrating Operational Feasibility," July 2001, Appendix 1.
[29] Census Bureau Standard: Definitions for Survey and Census Metadata.
[30] Ibid.
[31] Census Bureau, "Meeting 21st Century Demographic Data Needs-Implementing the American Community Survey: July 2001, Report 1: Demonstrating Operational Feasibility," July 2001, Appendix 1.

ACS Development Program. A reference to the full set of testing, research, and development program activities that started in 1994 and will continue until the ACS is fully implemented in 2003.

ACS test sites. The ACS development program expanded from an initial four test sites to 31 test sites, comprising 36 counties. When the term ACS test site is used, it refers to data from these sites.

Allocation. Refers to imputation required when values for missing 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 persons in the household, other responding households, or persons believed to have similar characteristics. Such donors are reflected in a table referred to as an allocation matrix. *See also Assignments and Imputation.*

Assignment. Imputation method in which values for a missing item or an inconsistent item can be derived from the sample housing unit or person record. For example, a first name is often used to determine the sex of a person.

Coeffcient of variation (CV).[32] The relative standard error of an estimator; that is, the standard error as a proportion of the magnitude of the population value being estimated. This quantity may be estimated by substituting estimates for the standard error and population value of interest.

Computer Assisted Telephone Interviewing (CATI).[33] Method or mode of data collection using telephone interviews in which the questions (to be asked) are displayed on a computer screen and responses are entered directly into the computer. *See also Follow-up*.

Computer Assisted Personal Interviewing (CAPI).[34] Method or mode of data collection consisting of the interviewer asking questions displayed on a portable computer screen and entering the answers directly into the computer. *See also Follow-up.*

Current Population Survey (CPS).[35] Monthly sample survey of the U.S. population that provides employment and unemployment figures as well as current data about other social and economic characteristics of the population. Collected for the Bureau of Labor Statistics by the Census Bureau.

Edit Followup. Quality assurance activity of mail response records intended to identify missing or inconsistent responses. Forms failing an automated coverage and content edit are followed up by telephone.

---

[32] U.S. Census Bureau, Decennial Management Division Glossary, 1999.
[33] U.S. Census Bureau, Decennial Management Division Glossary, 1999.
[34] Johnson, Margo J. Anderson, Editor in Chief, Congressional Quarterly Press Staff, "Encyclopedia of the U.S. Census", 2000, Congressional Quarterly Inc., p407.
[35] Ibid.

Housing unit.[36] A house, apartment, mobile home, or trailer, group of rooms or single room occupied as a separate living quarters or if vacant, intended for occupancy as a separate living quarters. The definition of separate living quarters for Census 2000 is that the occupants live separately from any other individuals in the building and have direct access from outside the building or through a common hall. Additional criteria, such as the presence of a kitchen or cooking equipment for the exclusive use of the occupants, were used to define a housing unit in previous censuses.

Imputation.[37] Process used to resolve problems of missing, invalid, or inconsistent responses identified during editing. Responses or missing values on the edited record are changed to ensure that a plausible, internally coherent record is created. *See also Allocation and Assignment.*

Long form.[38] The decennial census questionnaire containing 100-percent (short form) and sample questions. Sent to a sample of addresses in the census, long forms typically contain the short-form person and housing items that all households are asked to provide. Whereas short-form items are generally limited to basic demographic and housing questions, long-form items cover such topics as income, employment, veteran status, transportation to work, education, and others. *See Short Form.*

Mailout-mailback.[39] Descriptive of the enumeration method in which the Postal Service delivers census and survey questionnaires to specific addresses and the respondents mail them back to the census takers, district office, headquarters, or processing office for processing. Mailout-mailback is a primary method of data collection for censuses and surveys today.

Master Address File (MAF).[40] The Census Bureau's permanent list of addresses for individual living quarters that is linked to the TIGER data base. *See also Topologically Integrated Geographic Encoding and Referencing (TIGER) System.*

Measurement error. Error when the response received differs from the "true" value due to the respondent, the interviewer, the questionnaire, the mode of collection, or the respondent's record-keeping system(s).

Nonresponse error.[41] Error caused by survey failure to get a response to one or possibly all of the questions. Indirect measures include the detail disposition rates (unweighted and weighted) of all the selected sample cases during data collection. Direct measures may require nonresponse follow-up.

---

[36] Ibid.
[37] U.S. Census Bureau, Decennial Management Division Glossary, 1999.
[38] Johnson, Margo J. Anderson, Editor in Chief, Congressional Quarterly Press Staff, "Encyclopedia of the U.S. Census", 2000, Congressional Quarterly Inc., p407.
[39] Ibid.
[40] Ibid.
[41] Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys," Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, July 2001

Nonsampling error. Generally means all error (e.g., coverage, nonresponse, interviewer, respondent, instrument, mode) that does not arise merely because a sample of the population is measured.

Overcoverage. [42] Extent to which a frame includes more elements than the target population, including duplicate elements.

Processing error. Error during data editing, coding, capture (keying and scanning), imputation, and tabulation.

Relevance.[43] The qualitative assessment of the value contributed by the data. Value is characterized by the degree to which the data serve to address the purposes for which they are produced and sought by users (including mandate of the agency, legislated requirements, etc.)

Sampling frame.[44] Any list or device that, for purposes of sampling, delimits, identifies, and allows access to the sampling units, which contain elements of the sampled population. The frame may be a listing of persons, housing units, businesses, records, land segments, etc. One sampling frame or a combination of frames may be used to cover the entire sampled population.

Sampling error.[45] The difference between a sample result and the result from a complete count taken under the same conditions.

Short form.[46] The decennial census questionnaire requesting basic demographic and housing information. *See also Long form.*

Stratum, strata.[47] A sampling stratum is a grouping or classification that has a similar set of characteristics based on the previous census.

Subsampling. Refers to the sampling of a sample. In the C2SS, the cases that are not completed by mail or through a telephone interview become eligible for CAPI interviewing. However, only one out of three of these cases are actually interviewed. This winnowing of the sample is referred to as subsampling.

Survey quality. [48] The elements of quality consist of the relevance, accuracy, timeliness, accessibility, interpretability, and coherence of the data.

---

[42] Census Bureau Standard: Definitions for Survey and Census Methods
[43] Ibid.
[44] Ibid.
[45] U.S. Census Bureau, Decennial Management Division Glossary, 1999
[46] Johnson, Margo J. Anderson, Editor in Chief, Congressional Quarterly Press Staff, "Encyclopedia of the U.S. Census", 2000, Congressional Quarterly Inc., p407.
[47] U.S. Census Bureau, Decennial Management Division Glossary, 1999
[48] Census Bureau Standard: Definitions for Survey and Census Metadata

Operational Response Rates. Measure operational success in obtaining a response from a eligible unit. In the C2SS, operational response rates were calculated for the mail, CATI, and CAPI operations. *See also CATI and CAPI.*

Survey Response Rates. Measure total response across all three modes of data collection for the survey. In the C2SS, the survey response rate is weighted to reflect the sample design, including the CAPI subsampling.

Telephone Questionnaire Assistance (TQA). Telephone interviewers who field a spectrum of general and content-related survey questions in providing assistance to households in sample.

Timeliness. [49] Length of time between data availability and the event or phenomenon it describes (context of value and use).

Topologically Integrated Geographic Encoding and Referencing (TIGER) system.[50] A computer database that contains all census-required map features and attributes for the United States and its possessions, plus the specifications, procedures, computer programs, and related input materials required to build and use it.

Tract.[51] Small, relatively permanent statistical subdivisions of counties delineated by local committees of census data users in accordance with Census Bureau guidelines for the purpose of collecting and presenting decennial census data. These neighborhoods contain between 1,000 and 8,000 people, typically approximately 1,700 housing units and 4,000 people. Tracts are designed to have homogeneous population characteristics, economic status, and living conditions at the time they are established. Census tract boundaries normally follow visible features but may follow governmental unit boundaries and other nonvisible features.

Undercoverage. [52] The extent to which a frame includes fewer than the sampled population.

Undercount.[53] The total number of people missed in the census. The difference between the overcount and the undercount is the net undercount.

Weighting. A series of survey adjustments. Survey data are traditionally weighted to adjust for the sample design, the effects of nonresponse, and to correct for survey undercoverage error. *See also Nonresponse and Undercoverage.*

---

[49] Census Bureau Standard: Definitions for Survey and Census Methods
[50] Ibid.
[51] Ibid.
[52] Census Bureau Standard: Definitions for Survey and Census Methods
[53] Johnson, Margo J. Anderson, Editor in Chief, Congressional Quarterly Press Staff, "Encyclopedia of the U.S. Census", 2000, Congressional Quarterly Inc., p407.