

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

Issued December 2003

Report 3: Testing the Use of Voluntary Methods



AMERICAN
COMMUNITY
SURVEY

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

Helping You Make Informed Decisions

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

United States
**Census
2010**

Acknowledgments



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



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



U.S. CENSUS BUREAU
Charles Louis Kincannon,
Director

Hermann Habermann,
Deputy Director and
Chief Operating Officer

Deborah H. Griffin, David A. Raglin, Theresa F. Leslie, Pamela D. McGovern, and Joan K. Broadwater were the key authors of this report.

Other individuals who contributed to the design and implementation of the test and production of this report include:

Mark E. Asiala
Nancy A. Bates
Chester E. Bowie
Fern J. Bradshaw
Graham A. DeJong
Theresa J. DeMaio
Barbara N. Diskin
Karen C. Field
Donald P. Fischer
Nancy M. Gordon
David L. Hubble
Debra L.U. Klein
Ashley D. Landreth
Gail A. Leithauser
Elizabeth A. Martin
Lawrence S. McGinn
Grace T. Moe
Alfredo Navarro
Robert J. Packard
Kathleen M. Styles
Jefferson D. Taylor
Anthony G. Tersine, Jr.
Alan R. Tupek
Preston J. Waite

TABLE OF CONTENTS

EXECUTIVE SUMMARY	v
1. INTRODUCTION	1
2. BACKGROUND	1
3. METHODOLOGY	2
3.1 Sample Design and Experimental Treatments	2
3.2 Preparatory Materials	4
3.3 Quality and Performance Measures	4
3.4 Cost and Workload Projections	5
4. RESULTS	5
4.1 Respondent Cooperation	6
4.1.1 What was the effect on mail cooperation?	6
4.1.2 What was the effect on telephone cooperation?	7
4.1.3 What was the effect on personal visit cooperation?	8
4.1.4 Was there an effect on how ACS data were collected?	8
4.1.5 Was there an effect, by race or ethnicity, on how ACS data were collected? ...	10
4.2 Data Quality	11
4.2.1 What impact did voluntary methods have on the percentage of the initial sample that was interviewed?	11
4.2.2 What do we estimate is the impact of voluntary methods on the reliability of ACS estimates?	12
4.2.3 What additional impact does the use of voluntary methods have on the reliability of ACS estimates for small population groups?	13
4.2.4 What was the effect on unit nonresponse?	13
4.2.5 What was the effect on item nonresponse?	14
4.2.6 Was there an effect, by race or ethnicity, on item nonresponse?	15
4.3 Costs and Workloads	16
4.3.1 What are the projected changes in operational workloads?	16
4.3.2 What are the projected increases in costs for a change to voluntary methods? ..	16
5. REFERENCES	17

APPENDICES

- Appendix 1: Summary of Voluntary and Mandatory Differences in Mail Materials
- Appendix 2: Detailed Computation of Cooperation Rates, Weighted Survey Response Rates, Interview Mode Distributions, and Interview Rates
- Appendix 3: Detailed Computation of Data Completeness Rates
- Appendix 4: Derivation of Revised Sample Size
- Appendix 5: Workload Definitions and Determination
- Appendix 6: Cost Model

TABLES

Table 1. Distribution of Sample Across High and Low Response Strata	3
Table 2. Mail Cooperation Rates (2003 compared with 2002)	6
Table 3. Mail Cooperation Rates (2003 compared with 2003)	7
Table 4. Telephone Cooperation Rates (2003 compared with 2002)	7
Table 5. Personal Visit Cooperation Rates (2003 compared with 2002)	8
Table 6. Interview Mode Distributions (2003 compared with 2002)	9
Table 7. Interview Mode Distributions, by Race and Ethnicity (2003 compared with 2002) ..	10
Table 8. Survey Interview Rates (2003 compared with 2002)	12
Table 9. Total Number of Interviews	13
Table 10. Survey Response Rates (2003 compared with 2002)	14
Table 11. Data Completeness Rates (2003 compared with 2002)	15
Table 12. Data Completeness Rates, by Race and Ethnicity (2003 compared with 2002)	15
Table 13. Comparison of Operational Workloads	16
Table 14. Projected Survey Costs	17

EXECUTIVE SUMMARY

At the request of Congress, federal agencies, and other data users, the Census Bureau conducted research and produced two reports to address specific questions about operational feasibility and survey quality of the American Community Survey (ACS). More recently, issues of questionnaire content and increasing public privacy concerns led the Congress to ask if the ACS could be implemented as a voluntary, rather than a mandatory, survey.

Working closely with staff of the Technology, Information Policy, Intergovernmental Relations and Census Subcommittee and the House Government Reform Committee, a test was designed to provide answers to key questions about the impact, if any, that a change to voluntary methods would have on mail response, survey quality, and costs. This report summarizes the results of this test that was conducted between March and June of 2003. The test was implemented as designed with no problems encountered to bring any of the findings into question. The sample was sufficient to draw the following major conclusions:

A dramatic decrease occurred in mail response when the survey was voluntary (see page 6). The mail cooperation rate fell by over 20 percentage points. This decline has important consequences. The design of the ACS relies on high rates of response by mail and telephone for two reasons. First, personal visit follow-up activities are significantly more expensive than mail and telephone data collection activities (about ten times as great). Second, the nonrespondents after mail and telephone attempts are subsampled for personal visit follow-up. It is for these very reasons that a large drop in the portion of the sample that is interviewed by mail or telephone follow-up results in a concurrent drop in reliability and an increase in costs.

The reliability of estimates was adversely impacted by the reduction in the total number of completed interviews (see page 11). The ACS uses a two-phase sample design. Under this design only a subsample of nonrespondents after mail and telephone are eligible for personal visit follow-up. To produce reliable estimates with an initial sample of 3 million, about 70 percent of the initial sample needs to be interviewed. For a 3 million initial sample, this would result in roughly 2.2 million completed interviews. We monitor interview rates, the ratio of final interviews to the initial sample, to measure the impact of mail and telephone response on the sample size and thus, on the reliability of ACS estimates. In this test, about 72 percent of the initial sample was interviewed using mandatory methods. Using voluntary methods, this rate fell to about 60 percent. Voluntary methods (with an interview rate of 60 percent) would not provide the required 2.2 million interviews unless the initial sample size was increased.

The decrease in cooperation across all three modes of data collection resulted in a noteworthy, but not critical, drop in the weighted survey response rate (see page 13). Despite increased personal visit follow-up workloads and greater levels of refusals by mail and telephone, personal visit interviewers were successful in obtaining high levels of response. We produce weighted survey response rates to isolate the potential impact of nonresponse on survey estimates. Weighted survey response rates are calculated as the ratio of completed interviews

from mail, telephone, and personal visit follow-up to the universe of cases eligible to be interviewed. The weighted survey response rate dropped from nearly 98 percent (using mandatory methods) to about 93 percent (using voluntary methods).

The estimated annual cost of implementing the ACS would increase by at least \$59.2 million if the survey was voluntary and reliability was maintained (see page 16). We estimate that the initial sample size for a voluntary ACS with a 60 percent interview rate would need to be increased to about 3.7 million addresses annually to maintain reliability. We project that the associated costs would increase by at least \$59.2 million a year. Table A compares projections of the initial sample sizes, interview rates, completed interviews, and total costs for an ACS conducted using mandatory versus voluntary methods.

Table A. Comparison of initial sample sizes, interview rates, completed interviews, and total costs for a voluntary and a mandatory ACS

	Initial sample	Interview rate	Completed interviews	Total Cost
Voluntary	3.7 million	60 percent	2.2 million	\$214 million
Mandatory	3.0 million	72 percent	2.2 million	\$155 million
Difference	0.7 million increase	12 percentage point decrease	maintained	\$59 million increase

Levels of item nonresponse for the data collected under voluntary and mandatory methods were very similar (see page 14). Respondents provided answers to over 94 percent of the required questions under both voluntary and mandatory data collection methods. We believe that this indicates that once persons decided to complete the interview, they were equally willing to provide the requested information.

The use of voluntary methods had a negative impact on traditionally low response areas, that will compromise our ability to produce reliable data for these areas and for small population groups such as Blacks, Hispanics, Asians, American Indians, and Alaska Natives (see page 13). The current ACS design, based on mandatory methods, results in reduced levels of reliability for some areas and population groups due to low levels of mail response. This is a direct consequence of the additional number of cases that are subject to subsampling. In a voluntary setting, the problem is magnified. The proportion of completed interviews after mail and telephone attempts decreased to less than 50 percent for Black and Hispanic households. This has a significant impact on the total number of interviews and thus, on the reliability of estimates for these population groups.

The change to voluntary methods had the greatest impact on areas that have traditionally high levels of cooperation and on White and non-Hispanic households (see pages 9 and 10). The test results show a shift of 13 percentage points from mail and telephone to personal visit in areas with traditionally high levels of response. In areas with traditionally lower levels of

cooperation, the shift was 10 percentage points. The relative increase in the sample collected by personal visit was greatest for Whites and non-Hispanics (50 percent and 45 percent, respectively). The relative increase was less than 20 percent for Blacks, Hispanics, and All Other Races.

1. INTRODUCTION

To meet the challenges of rapid demographic and technological changes and the needs of stakeholders, the Census Bureau developed the American Community Survey (ACS). The ACS is one of three program components required to achieve the 2010 reengineering strategy goals. The current design of the ACS, with a 3 million sample, is expected to produce data of comparable quality to the sample data produced in the decennial census. Since the Census Bureau designed the ACS to replace the mandatory decennial census long form, we have conducted it as a mandatory survey since its inception.

In Fall 2002, Congress requested that the Census Bureau conduct a test to assess the effects of a voluntary ACS on mail response rates and associated follow-up costs (Weldon and Miller, 2002). Revised mail materials were designed and tested and telephone and personal visit interviewing procedures were modified. We printed and released all new materials to begin this test in March 2003. The test was implemented in March through June of 2003. The ACS has since converted back to the use of mandatory methods.

This report provides a description of the test design and evaluation methodology. Key results are summarized on the impact of the shift to voluntary data collection methods on (1) respondent cooperation; (2) data quality; and (3) workloads and costs. The voluntary results from the March and April 2003 sample panels are compared to the mandatory ACS results from the March and April 2002 sample panels. Refer to the Methodology Section for a more complete description.

This report includes survey cost projections for conducting the ACS under the current design using mandatory methods at full implementation levels (yearly sample of 3 million addresses). We premise projected survey costs for conducting the ACS on a voluntary basis on two assumptions. The first premise holds the sample size constant and assumes that accuracy would decrease as a consequence of reductions in mail response. This loss in reliability is considered another cost of a change to voluntary methods. The second premise is based on a larger initial sample size that we would need to maintain the same levels of reliability as under the current design.

This report was prepared on an expedited basis to meet Congressional needs. The Census Bureau plans to issue a second report with additional analysis of the test results. That report will include detailed analysis of the other mandatory and voluntary treatments that we included as part of the test.

2. BACKGROUND

The Census Bureau conducts the ACS continuously on independent monthly samples of addresses. Since 2000, a sample of about 70,000 addresses in 1,239 counties nationwide has been selected each month. In full implementation, this monthly sample size would increase to

approximately 250,000 addresses spread across all 3,141 counties. The data for each sample are collected over a three-month period. The design of the ACS relies on optimizing three modes of data collection—mail, telephone follow-up, and personal visit follow-up. Significant cost implications apply to each mode; personal visit follow-up is by far the most expensive mode of data collection.

The Census Bureau first attempts collecting all ACS data using mailout/mailback techniques. To maximize mail response, several mailings are used including an advance letter, reminder card, and a targeted second mailing to nonrespondents. Call centers provide telephone assistance to help households complete the forms they receive in the mail. Mail returned forms are data captured by keying and reviewed for completeness. Telephone interviewers resolve incomplete forms in an edit follow-up operation.

ACS interviewers follow up on nonrespondents to the mailout through separate telephone and personal visit follow-up operations. For example, data collection for the March 2003 panel started when we mailed the survey questionnaires in late February. In April, interviewers conducted a telephone follow-up operation to collect data for nonresponding addresses for which a telephone number was available. At the end of April, a sample of 1-in-3 of the addresses that did not respond by mail or telephone was selected for a personal visit follow-up operation conducted in May. A 2-in-3 sample of addresses that could not be mailed (due to incomplete address information) was added to the personal visit follow-up workload.

This combination of data collection activities has been very successful. In 2001 and 2002, the weighted survey response rates were 96.7 percent and 97.7 percent, respectively and more than half of the addresses had mailed back completed forms. See Appendix 2 for details on how weighted survey response rates are calculated.

3. METHODOLOGY

3.1 Sample Design and Experimental Treatments

The ACS sample for 2003 includes twelve monthly samples of about 70,000 addresses for a total of about 840,000 addresses. The ACS sample cases for March and April 2003, about 140,000 addresses, formed the universe for the ACS Voluntary Test.

We designed the sample for the ACS Voluntary Test to study four experimental mail treatments—two mandatory and two voluntary. One mandatory mail treatment was identical to the mail treatment used in prior years and provided a control to previous years; the other tested mandatory mail materials that we had revised with an eye toward increased user-friendliness. The Census Bureau will include detailed analysis of these two mandatory treatments in a future report. In this report, the 2002 Current Mandatory approach is the benchmark mandatory treatment. We tested two voluntary mail treatments. One used a standard survey approach to explain the voluntary nature of the survey, the approach that the Census Bureau uses for its

current surveys. That treatment, called the Standard Voluntary treatment, is the focus of this report. A second voluntary treatment, which we will analyze in a future report, explained more directly that the survey was voluntary. Since the primary focus of this test was to evaluate the effect of voluntary methods on the ACS, we evenly distributed 75 percent of the combined 2003 March/April sample to the two voluntary mail treatments and the remaining 25 percent between the two mandatory mail treatments.

Unlike the mail portion, the test used only voluntary methods in the telephone and personal visit follow-up operations. We concluded that assigning both voluntary and mandatory methods to a sample of cases or a sample of interviewers introduced potential implementation risks to the study that the use of one method could avoid during the test period. Without a mandatory control for follow-up operations in 2003, this test will compare 2003 results with 2002 results that were conducted using mandatory methods. For all analyses of the effect of voluntary methods on overall survey quality, costs, and workloads, we used only sample cases that were voluntary in all phases of data collection.

The sample design divided the universe into two strata, high response areas (HRA) and low response areas (LRA). We created these strata using tract-level long form mail return rates from Census 2000. Based on data from the 2001 ACS, people in the LRA stratum are younger, more Hispanic and non-White, and have more other relative and non relative household members than people in the HRA stratum. The LRA stratum also has less people with college educations, more renters, more households who speak a language other than English at home, and more households with lower incomes compared with the HRA stratum. Within strata, experimental treatments were designated in a systematic manner to ensure that we assigned 75 percent of the sample evenly to the two voluntary treatments and 25 percent of the sample to the two mandatory treatments. (Asiala, 2003 and Tersine, 2003 include more details of the stratification and sample selection.)

Table 1 documents the sample for the 2003 combined March/April treatments. We applied this stratification to the 2002 sample and we show the results in Table 1. The shaded portion of the table shows the sample for which the Standard Voluntary message was used in the mailout, followed-up by voluntary telephone and personal visit operations. This sample is half of the sample allocated to the voluntary treatments or about 37 percent of the full sample and is used throughout this report to answer most of the study questions. We plan to use the remaining sample to answer additional questions in a second report.

Table 1. Distribution of Sample Across High and Low Response Strata

Stratum	March/April 2002 Mandatory	March/April 2003 Standard Voluntary Treatment	March/April 2003 Remaining Experimental Treatments
Overall	134,632	51,712	86,187
High Response Areas	100,933	38,740	64,567
Low Response Areas	33,699	12,972	21,620

3.2 Preparatory Materials

As mentioned, four different sets of mail treatment materials were tested—two mandatory and two voluntary. Copies of the final mail treatment materials are available (Leslie, 2003). Before we fielded these materials, we cognitively tested them. Landreth (2003) documents the testing protocol, complete findings, and recommendations. Appendix 1 of this report highlights the key messages used to convey that the survey was voluntary in the 2003 Standard Voluntary mail materials. Also, shown are comparable messages for the 2002 Current Mandatory mail materials.

All telephone and personal visit follow-up operations in March through June of 2003 were conducted using voluntary methods, regardless of mail treatment. To make this change, staff revised the automated data collection instrument scripts to remove all references to mandatory data collection and to provide appropriate information on the survey now being voluntary. Call center and field interviewers received detailed memoranda explaining the change to voluntary, including updated pages to their interviewer manuals. In addition, staff updated the brochure, used in the field to gain cooperation, to remove references to mandatory data collection. These procedures required the interviewer to explain that the survey was voluntary if the respondent had not received the mail materials. We did not tell the telephone or personal visit interviewers that this was a test.

3.3 Quality and Performance Measures

The primary comparisons for this test are based on combined sample data from March and April 2003 compared to combined sample data from March and April 2002. We calculated all quality and performance measures nationally and at the stratum level. They include:

- Mail, Telephone, and Personal Visit Cooperation Rates - measures of respondent behavior in each of the three data collection modes,
- Interview Mode Distributions - measures of how the ACS interviewed occupied households across these modes,
- Interview Rates - measures the impact of subsampling and nonresponse on the final number of completed interviews.
- Weighted Survey Response Rates - measures of unit nonresponse,
- Data Completeness Rates - measures of item nonresponse, and
- Total Number of Interviews- measures used to estimate reliability.

Interview mode distributions and data completeness measures also are assessed by race and ethnicity of the first household member. We weighted the performance and quality measures by their final probabilities of selection, including subsampling factors, except for the interview rate, which is weighted by the probability of selection before subsampling. Appendices 2 and 3 include details on how each of these measures were calculated. Refer to Raglin (2003) and McGovern (2003) for the detailed specifications.

Most of the tables in the results section compare the estimated quality and performance measures for the 2003 Standard Voluntary and 2002 Current Mandatory treatments. One table compares two 2003 treatments. We show percentage point differences between the two measures. The tables round the rates and differences to one decimal place. Due to rounding, the difference column may not always reflect the exact difference between the two displayed estimates. The tables include the margins of error of each difference, indicating the 90 percent confidence interval around the difference. The 90 percent confidence interval tells us that if all possible samples under the sample design were selected independently and surveyed under the same conditions, approximately 90 percent would fall within the range of the estimates provided. We calculated direct estimates of the standard errors for all estimates in this report using standard ACS variance estimation methods. The table notes those differences determined to be statistically significant at the 90 percent confidence level. In addition, we refer in the text to relative differences which provide a measure of the change relative to the previous year's estimate. These rates were calculated as the ratio of the difference between the 2003 and 2002 estimates to the 2002 estimate.

3.4 Cost and Workload Projections

This report includes cost and workload projections for full implementation (a monthly sample of 250,000 addresses or an annual sample of 3 million addresses) and for a larger sample that would maintain the reliability of the current design. We used the 3 million sample size and key operational parameters to model the workloads and costs of the mail, telephone, and personal visit follow-up operations. The parameters for the mandatory approach were based on data from March through June of 2002, under the conditions that the survey currently is being conducted. We have based the voluntary parameters on the March through June 2003 test results. To produce annual projections, the parameters (which were based on data from March through June) were seasonally adjusted. We discuss these parameters, the workloads, and the costs in Appendices 4, 5 and 6.

The workload models allowed us to estimate the final sample sizes and therefore to assess the loss in reliability that would occur under the voluntary option. From those data we determined the increase in the sample that we would need to maintain the same reliability of ACS estimates. We ran the revised initial sample through the voluntary model to project the workloads and costs of a voluntary ACS that met the ACS program's minimum reliability standards.

4. RESULTS

This section answers key questions on respondent behavior, data quality, costs, and workloads. Unless otherwise noted, "mandatory" refers to the March/April 2002 mandatory results (i.e., the Current Mandatory methods) and "voluntary" refers to the results for the March/April 2003 sample cases that used the Standard Voluntary treatment in the mail and voluntary methods in telephone and personal visit follow-up.

4.1 Respondent Cooperation

Cooperation rates are the best measure of respondent behavior because they measure the rate of response for only those persons or households who are contacted. The universe for calculating cooperation rates is the subset of occupied units that were contacted. In this report, we calculated cooperation rates for each mode of data collection—mail, telephone, and personal visit. We use these data to answer questions about whether the change to voluntary methods had an impact on the behavior of the public in choosing to participate in the ACS.

4.1.1 What was the effect on mail cooperation?

The change to voluntary methods resulted in a very large drop in mail cooperation, as shown in Table 2. The difference in the mail cooperation rates was 20.7 percentage points nationally, 22.1 percentage points in high response areas, and 16.0 percentage points in low response areas. High response areas show a greater percentage point drop than low response areas suggesting that areas with fewer minorities, higher levels of education, and higher household incomes were more likely to be affected by the change to voluntary. However, both areas had similar relative decreases - a 34.3 percent drop in high response areas and a 36.6 percent drop in low response areas.

Table 2. Mail Cooperation Rates (2003 compared with 2002)

Stratum	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	59.5	38.8	-20.7	± 0.6	Yes
High Response Areas	64.6	42.4	-22.1	± 0.7	Yes
Low Response Areas	43.6	27.7	-16.0	± 1.1	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 eligible mail workload. These cases received the Current Mandatory treatment.
2003 Voluntary results are based on the March and April 2003 eligible mail workload designated for the Standard Voluntary mail treatment.

Since the voluntary test included a mandatory mail control treatment, we also compared the March/April cooperation rates for the 2003 mandatory and 2003 voluntary treatments. Table 3 shows that these results were very similar. The combined change in messages on the envelope (dropping the mandatory message) and to the other mail materials (explaining that the survey was voluntary) clearly affected respondent behavior. Tables 2 and 3 also show a slight decline in mail cooperation from 2002 to 2003 (59.5 versus 57.3).

Table 3. Mail Cooperation Rates (2003 compared with 2003)

Stratum	2003 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (Voluntary - Mandatory) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	57.3	38.8	-18.5	± 1.2	Yes
High Response Areas	61.5	42.4	-19.0	± 1.4	Yes
Low Response Areas	43.9	27.7	-16.2	± 2.3	Yes

KEY: 2003 Mandatory results are based on the March and April 2003 eligible mail workload designated for the Current Mandatory mail treatment.
2003 Voluntary results are based on the March and April 2003 eligible mail workload designated for the Standard Voluntary mail treatment

4.1.2 What was the effect on telephone cooperation?

Taking the survey on a voluntary basis also resulted in lower levels of telephone cooperation as shown in Table 4. After mail data collection and before selecting the sample for personal visit follow-up, the ACS conducts a telephone follow-up operation. Due to the large drop in mail cooperation, the workload for telephone follow-up increased. The universe eligible for telephone interviewing therefore was different for the voluntary treatment group—it included more households that chose not to respond because they had acknowledged that the survey was voluntary.

Comparing the combined March/April telephone cooperation rates for the 2003 voluntary treatment with the combined March/April telephone cooperation rates in 2002, the overall drop was 14.2 percentage points. The 2003 rates are based only on the sample cases that received the Standard Voluntary mail materials and were eligible for telephone follow-up. In high and low response areas, the decrease in telephone cooperation was 15.5 percentage points and 10.2 percentage points, respectively. The relative drop in high response areas was 18.9 percent, over 5 percentage points higher than the relative drop in low response areas. This suggests that a greater impact was found in areas that are traditionally more cooperative. The use of voluntary methods negatively affected respondent behavior to a request for a telephone interview.

Table 4. Telephone Cooperation Rates (2003 compared with 2002)

Stratum	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	80.7	66.5	-14.2	± 1.6	Yes
High Response Areas	82.1	66.5	-15.5	± 1.8	Yes
Low Response Areas	76.4	66.2	-10.2	± 3.3	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 eligible telephone follow-up workload. These cases received the Current Mandatory mail materials and were followed up by telephone using mandatory methods.
2003 Voluntary results are based on the March and April 2003 eligible telephone follow-up workload designated for the Standard Voluntary mail treatment and followed up by telephone using voluntary methods.

4.1.3 What was the effect on personal visit cooperation?

The personal visit follow-up operation made up some, but not all, of the reduced cooperation in the mail and telephone phases, as shown in Table 5. As was true in telephone follow-up, the reduction in cooperation in the mail and telephone stages resulted in a much larger workload for personal visit follow-up. A greater proportion of sample cases were households that refused to participate by mail and by telephone. The universe for voluntary personal visit follow-up in 2003 therefore was quite different from the universe for a mandatory personal visit follow-up in 2002. The 2003 voluntary data are based on cases designated for the Standard Voluntary mail treatment for which data were not collected by either mail or telephone follow-up and that were selected for personal visit follow-up.

Comparing the combined March/April personal visit cooperation rates for the 2003 voluntary treatment with the combined March/April personal visit cooperation rates for 2002 shows an overall drop of 6.7 percentage points. Differences of 7.5 percentage points and 4.9 percentage points were found in high and low response areas, respectively. The relative drop in high response areas was 7.8 percent, while the relative drop in low response areas was 5.1 percent. Here again, we see the greater impact occurring in high response areas. Although personal visit cooperation rates remained high, knowing that the survey was voluntary affected respondent behavior to a personal visit request to complete the survey.

Table 5. Personal Visit Cooperation Rates (2003 compared with 2002)

Stratum	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	95.6	89.0	-6.7	± 0.8	Yes
High Response Areas	95.7	88.2	-7.5	± 1.0	Yes
Low Response Areas	95.6	90.7	-4.9	± 1.3	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 eligible personal visit follow-up workload. These cases received the Current Mandatory mail treatments and were followed up by personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 eligible personal visit workload designated for the Standard Voluntary mail treatment and followed up by personal visit using voluntary methods.

4.1.4 Was there an effect on how ACS data were collected?

The change to voluntary methods resulted in a shift of respondents from participating by mail to participating by telephone or personal visit follow-up. Table 6 provides an illustration of how data were collected when mandatory versus voluntary methods were used. Interview mode distributions answer questions about whether the change to voluntary had an impact on the behavior of the public in choosing **when** to participate in the ACS. Interview mode distributions are based on the final weighted estimate of occupied units and indicate the proportion of the data that were collected by mail, telephone, and personal visit. These distributions were weighted to reflect the probabilities of selection and the subsampling factors and do not summarize the percent of actual interviews collected by each mode. (We include those results in Table 9.) The

distributions were produced overall, by strata, and by race and Hispanic origin of the first household member.

There are two important reasons to look at these distributions—cost and reliability. Personal visit interviews are much more costly than mail or telephone interviews, so a drop in the proportion of interviews collected by mail and telephone will signal an increase in survey costs. In addition, because the ACS selects a subsample of nonrespondents after mail and telephone to go to personal visit follow-up, there are reliability concerns. If we see a drop in the percent of interviews collected by mail and telephone, we will have fewer total interviews and thus, a reduction in reliability.

Table 6. Interview Mode Distributions (2003 compared with 2002)

Stratum and Mode		2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Mail	Overall	58.9	40.5	-18.4	± 0.7	Yes
	High Response Areas	63.8	44.2	-19.5	± 0.8	Yes
	Low Response Areas	43.5	29.0	-14.6	± 1.2	Yes
Telephone	Overall	9.5	15.1	5.6	± 0.4	Yes
	High Response Areas	9.5	15.6	6.0	± 0.4	Yes
	Low Response Areas	9.4	13.8	4.4	± 0.7	Yes
Personal Visit	Overall	31.6	44.3	12.7	± 0.7	Yes
	High Response Areas	26.7	40.2	13.5	± 0.9	Yes
	Low Response Areas	47.1	57.2	10.2	± 1.4	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 sample designated for the Current Mandatory mail treatment and followed up by telephone and/or personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 sample designated for the Standard Voluntary mail treatment and followed up by telephone and/or personal visit using voluntary methods.

In 2002, using mandatory methods, mail accounted for 58.9 percent of all occupied interviews. When voluntary methods were used in 2003, the distribution shifted to 40.5 percent. Using mandatory methods, telephone accounted for 9.5 percent of occupied interviews as compared to 15.1 percent using voluntary methods. Therefore, mail and telephone combined for 68.4 percent of all occupied interviews using mandatory methods, but only 55.7 percent using voluntary methods.¹

The shifting of cases from mail or telephone to personal visit has an important impact on data quality. A greater number of cases remained as nonrespondents after mail and telephone attempts and therefore, a larger proportion of cases were subsampled and carried larger sampling weights. Under mandatory methods, since 68.4 percent of the interviews for occupied units were a result of mail and telephone efforts, 31.6 percent were personal visit cases with additional

¹The percent mail plus the percent telephone differs slightly from the percent mail and telephone combined due to rounding.

subsampling weights. Using voluntary methods, the 31.6 percent rose to 44.3 percent, a 12.7 percentage point increase in the proportion of interviews that have these higher weights.² Section 4.2.1 explains how this impacts reliability.

4.1.5 Was there an effect, by race or ethnicity, on how ACS data were collected?

Table 7 shows the interview mode distributions by race and ethnicity of the first household member. We see shifts from mail to telephone and personal visit across all race and ethnic groups. We see the greatest impact of voluntary methods in White and non-Hispanic households. Table 6 shows a national shift of 12.7 percentage points from mail or telephone interviews to personal visit interviews. That rate ranged from 14.0 percentage points for White households to 6.1 percentage points for All Other Races (includes American Indian and Alaska Natives, Asians, Native Hawaiian and Other Pacific Islanders, and Some Other Race). Hispanic and Black households had the same shifts, 8.4 percentage points.

Table 7. Interview Mode Distributions, by Race and Ethnicity (2003 compared with 2002)

Race/Ethnicity and Mode		2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Mail	White	63.3	43.4	-19.9	± 0.7	Yes
	Black	35.2	22.2	-13.1	± 1.5	Yes
	All Other Races	41.7	30.4	-11.3	± 2.0	Yes
	Hispanic	32.2	20.1	-12.2	± 1.8	Yes
	Not Hispanic	61.1	42.1	-19.1	± 0.7	Yes
Telephone	White	9.0	14.9	5.9	± 0.4	Yes
	Black	11.3	16.0	4.6	± 1.3	Yes
	All Other Races	12.9	18.0	5.1	± 1.3	Yes
	Hispanic	11.0	14.8	3.8	± 1.3	Yes
	Not Hispanic	9.5	15.4	5.9	± 0.4	Yes
Personal Visit	White	27.7	41.7	14.0	± 0.8	Yes
	Black	53.4	61.8	8.4	± 1.9	Yes
	All Other Races	45.5	51.6	6.1	± 2.4	Yes
	Hispanic	56.8	65.1	8.4	± 2.4	Yes
	Not Hispanic	29.4	42.6	13.2	± 0.8	Yes

Key: 2002 Mandatory results are based on the March and April 2002 sample designated for the Current Mandatory mail treatment and followed up by telephone and/or personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 sample designated for the Standard Voluntary mail treatment and followed up by telephone and/or personal visit using voluntary methods.

Note: The race and ethnicity of the household are based on the unedited response of person 1. If the Hispanic origin or race questions were not answered for person 1 in the household, the household is not included in the distribution. A household may be in more than one race category if person 1 reported more than one race, but a household cannot be both Hispanic and not Hispanic.

The relative increase in the sample collected by personal visit was greatest for Whites and non-Hispanics (50.5 percent and 44.9 percent, respectively). The relative increase was

²The percent mail plus the percent telephone differs slightly from the percent mail and telephone combined due to rounding.

15.8 percent for Blacks, 14.7 for Hispanics, and 13.5 percent for All Other Races. This indicates that the use of voluntary methods resulted in more of the sample for non-minority households shifting from telephone or mail collection to personal visit collection. It also indicates a greater relative loss in sample for these groups. However, it must be noted that the proportion of the sample based on personal visit data collection for Black, Hispanic, and All Other Race households was already quite low in 2002 using mandatory methods. The Census Bureau has been researching methods to address this differential by oversampling areas with low levels of mail response. The shift to voluntary brings these groups to an even lower level.

4.2 Data Quality

This section answers questions about the impact that a change to voluntary methods had on data quality. Data quality has four key dimensions - relevance, accessibility, timeliness, and accuracy (OMB 2001). This report focuses on accuracy implications. Accuracy is traditionally measured through the analysis of survey errors, including both 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.

Sampling error implications are measured by quantifying the effect that a reduction in mail response had on the final sample size and, thus, on the reliability of estimates. Nonsampling error is assessed by measuring both unit and item nonresponse. Unit nonresponse is the failure to obtain data from a unit and may occur because households are unwilling or unable to participate, or because an interviewer may be unable to make contact with a respondent (U.S. Census Bureau 2002). Item nonresponse occurs when a responding unit fails to provide complete and usable information for a particular item. Unit and item nonresponse were measured using survey response and data completeness rates.

An additional measure of data accuracy is needed to more completely describe the impact of voluntary methods. Survey response rates are defined for a specific, restricted universe and are used to measure the survey's ability to complete interviews with eligible sample units. Due to the use of subsampling, the universe of eligible sample units will vary as patterns of mail and telephone response vary. Cases that are subsampled out at the personal visit stage are no longer eligible sample units—they are represented by the cases selected to be in the personal visit subsample. To paint a clearer picture of overall data accuracy we produced interview rates. The interview rate is a combined measure of the impact of mail and telephone response (and therefore, subsampling) and personal visit response on the total number of completed interviews.

4.2.1 What impact did voluntary methods have on the percentage of the initial sample that was interviewed?

The ACS uses a two-phase sample design and therefore the final ACS sample is defined as the

combination of (1) cases interviewed by mail and telephone and (2) the subsample of nonresponse cases that are interviewed by personal visit. The interview rate is the ratio of the final number of completed interviews to the initial (or first phase) sample. Cases subsampled out in the personal visit stage are included in the denominator of the interview rate. Therefore, this rate is substantially lower than the survey response rate shown in Table 10 (on page 14). Interview rates are important to monitor as cooperation decreases because they provide a direct measure of the associated reduction in the final count of survey interviews. Appendix 2 provides additional details and definitions.

Table 8 shows that interview rates declined from 71.8 percent in 2002 (mandatory) to 60.2 percent in 2003 (voluntary). Significant differences were also seen in high and low response areas.

Table 8. Survey Interview Rates (2003 compared with 2002)

Survey Response Rate	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	71.8	60.2	-11.6	± 0.3	Yes
High Response Areas	75.4	62.9	-12.5	± 0.3	Yes
Low Response Areas	61.1	52.0	-9.1	± 0.6	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 sample designated for the Current Mandatory mail treatment and followed up by telephone and/or personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 sample designated for the Standard Voluntary mail treatment and followed up by telephone and/or personal visit using voluntary methods.

4.2.2 What do we estimate is the impact of voluntary methods on the reliability of ACS estimates?

The reliability expected in the current ACS design has been accepted by most data users. The ACS annual sample size of 3 million is designed to produce tract level estimates over a 5-year period. Alexander (2002) reported that the standard errors for a 5-year average estimate from the ACS will be about 1.33 times as large as the standard errors on a corresponding estimate from the census long form. To most data users, this decrease in reliability over the decennial long form sample is a worthwhile tradeoff in order to receive annual data. But, any reduction below the planned initial annual sample of 3 million is likely to reduce reliability to an unacceptable level. We have estimated that the variance of ACS estimates based on this reduced number of interviews would be about 23 percent higher than under the current design. The effect would likely vary by area. The Census Bureau believes that this deterioration of reliability would not allow the ACS to produce data of sufficient quality to replace the long form.

The ACS collects data by mail and telephone and then selects a subsample of the remaining nonrespondents for personal visit follow-up. Using mandatory methods, a majority of interviews are collected by mail and telephone. The drop in cooperation by mail and telephone when voluntary methods were used has a critical impact on the reliability of estimates for two reasons.

The major impact is a loss in the total number of interviews conducted. A second impact, however, is the change in the proportion of interviews that have higher weights due to the additional subsampling weights. As more cases shift to personal visit, a higher proportion of the completed interviews carry these higher weights. Both of these factors contribute to an increase in sampling error and thus, to a decrease in the reliability of sample estimates.

The reliability of ACS estimates would suffer if the survey were conducted using voluntary methods without an increase in the sample size. Table 9 summarizes the expected number of completed interviews resulting from the Current Mandatory methods and the tested voluntary methods assuming an initial annual sample of 3 million addresses.

Table 9. Total Number of Interviews

Completed Interviews	Current Reliability (Mandatory Methods)	Lower Reliability (Voluntary Methods)
Total	2,175,000	1,797,000
Mail	1,490,000	942,000
Telephone Follow-up	271,000	337,000
Personal Visit Follow-up	414,000	518,000

KEY: Current Reliability (Mandatory Methods) reflects 3 million initial annual sample and use of Current Mandatory methods.
Lower Reliability (Voluntary Methods) reflects 3 million initial annual sample and use of Standard Voluntary methods.

4.2.3 What additional impact does the use of voluntary methods have on the reliability of ACS estimates for small population groups?

As shown in Table 7 (on page 10), the percent of occupied interviews completed by mail for Blacks and Hispanics are lower than other race groups under the Current Mandatory design. This in turn results in higher proportions of the Black and Hispanic households being subject to the 1-in-3 subsampling for personal visit follow-up, than other households.

Table 7 shows that under voluntary methods the proportion of occupied interviews collected by mail fell to 22.2 percent for Black households and 20.1 percent for Hispanic households. All Other Race households fell to about 30.4 percent. The proportion of occupied interviews now represented by personal visit data rose to 61.8 percent for Black households, 65.1 percent for Hispanic households, and 51.6 percent for All Other Race households.

4.2.4 What was the effect on unit nonresponse?

The ACS experienced a decrease in the percent of completed interviews from eligible sample units when the survey was voluntary. Survey response rates are considered a key measure of survey quality and traditionally are produced for all surveys as an indicator of unit nonresponse. We calculated weighted survey response rates as the number of total interviews (complete and partial) across all three modes of data collection divided by the sum of the number of interviews (complete and partial), and the number of noninterviews (refusals, break-offs, non contacts, and

other noninterviews).³ We include only eligible sample housing units in the denominator of this rate. Appendix 2 provides details on how we computed these rates. The weighted survey response rate is higher than the interview rate (see Table 8, page 12) because only eligible cases in the final sample, after subsampling, are included in the denominator of the weighted survey response rate.

Table 10 compares the combined 2003 March/April weighted survey response rates to the combined March/April weighted survey response rates for 2002 to answer this question. The overall weighted survey response rate declined by 4.2 percentage points—from 97.6 percent to 93.4 percent. The high and low response areas show similar declines (4.3 percentage points and 3.9 percentage points, respectively). The relative drop in the weighted survey response rate was similar in high and low response areas (4.4 percent and 4.0 percent, respectively).

Table 10. Weighted Survey Response Rates (2003 compared with 2002)

Weighted Survey Response Rate	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	97.6	93.4	-4.2	± 0.4	Yes
High Response Areas	97.9	93.6	-4.3	± 0.4	Yes
Low Response Areas	96.8	92.9	-3.9	± 0.7	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 sample designated for the Current Mandatory mail treatment and followed up by telephone and/or personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 sample designated for the Standard Voluntary mail treatment and followed up by telephone and/or personal visit using voluntary methods.

4.2.5 What was the effect on item nonresponse?

In addition to studying unit nonresponse, it is important to assess the completeness of data at the item level. We calculated data completeness rates for all occupied housing units at the national and stratum levels and by race and ethnicity of the first household member. We define data completeness rates as the ratio of the total number of valid responses for all data items to the total number of data items requiring a response. (Refer to Appendix 3 for details of the calculation of the data completeness rates.) The data set used to calculate the completeness rates reflects the completeness of the data before they were edited. Consequently, the rates reflect the completeness of the returned questionnaires and interviews, which may differ slightly from the completeness of the final data.

Table 11 shows the data completeness rates overall and by strata for all housing and population questions. Comparisons of voluntary methods (2003) with mandatory methods (2002) show that

³To produce overall and strata level estimates of eligible sample units, the survey response rates were weighted by the final probabilities of selection, including subsampling factors.

the differences in the data completeness rates were low, although statistically significant overall and for the high response area stratum. These results suggest that, although differences in respondent behavior were found in their choice to participate in the interview, once the respondent made the choice to participate, the data provided were nearly as complete. In other words, under voluntary methods, there is no evidence that respondents who completed mail forms chose to leave many more questions blank nor that they provided fewer responses to telephone or personal visit interviewers.

Table 11. Data Completeness Rates (2003 compared with 2002)

Stratum	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
Overall	94.6	94.1	-0.5	± 0.4	Yes
High Response Areas	94.8	94.2	-0.6	± 0.4	Yes
Low Response Areas	94.0	94.0	0.0	± 0.6	No

KEY: 2002 Mandatory results are based on the March and April 2002 sample designated for the Current Mandatory mail treatment and followed up by telephone and/or personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 sample designated for the Standard Voluntary mail treatment and followed up by telephone and/or personal visit using voluntary methods.

4.2.6 Was there an effect, by race or ethnicity, on item nonresponse?

Table 12 shows these data completeness rates by race and ethnicity of the first household member. Again, even the statistically significant differences were small numerically. No significant differences were found on interviews collected from Black, All Other Races, or Hispanic households although significant differences were found for White and non-Hispanic households.

Table 12. Data Completeness Rates, by Race and Ethnicity (2003 compared with 2002)

Race/Ethnicity	2002 Mandatory (in percent)	2003 Voluntary (in percent)	Difference (2003 - 2002) (in percentage points)	Margin of Error of Difference (in percentage points)	Is the Difference Statistically Significant?
White	95.2	94.5	-0.7	± 0.4	Yes
Black	93.8	93.8	0.0	± 0.7	No
All Other Races	95.0	95.4	0.4	± 0.7	No
Hispanic	95.2	95.6	0.4	± 0.6	No
Not Hispanic	95.1	94.4	-0.7	± 0.4	Yes

KEY: 2002 Mandatory results are based on the March and April 2002 sample designated for the Current Mandatory mail treatment and followed up by telephone and/or personal visit using mandatory methods.
2003 Voluntary results are based on the March and April 2003 sample designated for the Standard Voluntary mail treatment and followed up by telephone and/or personal visit using voluntary methods.

Note: The race and ethnicity of the household are based on the unedited response of person 1. If the Hispanic origin or race questions were not answered for person 1 in the household, the household is not included in the distribution. A household may be in more than one race category if person 1 reported more than one race, but a household cannot be both Hispanic and not Hispanic.

4.3 Costs and Workloads

Based on the results described in section 4.2, an assessment was made of the increase in the initial sample that we would need to produce sample estimates with the same reliability as the Current Mandatory design. We would need to increase the initial annual sample of 3 million to approximately 3.7 million, as described in Appendix 4. This section provides projections of the workloads and costs of implementing the ACS under three designs - (1) an initial sample of 3 million using mandatory methods; (2) an initial sample of 3.7 million using voluntary methods (and achieving the same reliability); and (3) an initial sample of 3 million using voluntary methods (which would result in significantly less reliable estimates). We include workload and cost models in Appendices 5 and 6.

4.3.1 *What are the projected changes in operational workloads?*

Based on the results of the voluntary test, projections were made of the workloads for key data collection operations. Workloads include the full set of cases sent to each data collection operation and therefore include both completed interviews and noninterviews. Due to nonresponse, cases will, in many instances, be included in the workloads for more than one operation. For example, nonrespondents to mail and telephone may be included in the workload for mail, telephone, and personal visit follow-ups. Table 13 summarizes the projected annual workloads for mail, telephone, and personal visit follow-up for the current design and two alternative designs.

Table 13. Comparison of Operational Workloads

Operation	Current Reliability (Mandatory Methods)	Current Reliability (Voluntary Methods)	Lower Reliability (Voluntary Methods)
Mail	2,880,000	3,544,000	2,880,000
Telephone Follow-up	1,050,000	1,485,000	1,206,000
Personal Visit Follow-up	480,000	777,000	632,000

KEY: Current Reliability (Mandatory Methods) reflects 3 million initial annual sample and use of Current Mandatory methods.
 Current Reliability (Voluntary Methods) reflects 3.7 million initial annual sample and use of Standard Voluntary methods.
 Lower Reliability (Voluntary Methods) reflects 3 million initial annual sample and use of Standard Voluntary methods.

4.3.2 *What are the projected increases in costs for a change to voluntary methods?*

Table 14 uses the workloads in Table 13 to project the costs of the major data collection operations. All estimates are in FY 2005 dollars. The cost model assumes that the per unit cost of mail, telephone, and personal visit follow-up are consistent for both a voluntary and a mandatory survey. This test did not include Puerto Rico or group quarters and therefore direct estimates of workloads and costs were not available. Since we expect increased costs in Puerto Rico and for group quarters given a move to voluntary, we estimate that voluntary costs would increase by a similar factor. We believe that this is a reasonable estimate. Appendix 6 provides additional detail. We have included all other survey costs as fixed costs across options. These costs include headquarters staff costs to design, conduct, monitor, research and evaluate the ACS

operations. Included also are headquarters costs to process, tabulate, review, and disseminate ACS data products and to support, develop, and maintain the automated data collection instruments used for telephone and personal visit follow-up operations.

Table 14 indicates that conducting the ACS using voluntary methods on a sample of the necessary size to achieve the same reliability, would require at least an additional \$59.2 million per year. This estimate reflects direct data collection costs only and would not reflect potentially large one-time costs that the survey would incur for infrastructure changes, such as additional recruiting and hiring, training of additional staff, and the purchase of more equipment such as laptop computers.

Table 14. Projected Survey Costs

Operation	Current Reliability (Mandatory Methods)	Current Reliability (Voluntary Methods)	Lower Reliability (Voluntary Methods)
Mail	\$36,000,000	\$44,306,000	\$36,000,000
Telephone Follow-up	\$15,750,000	\$22,269,000	\$18,094,000
Personal Visit Follow-up	\$64,800,000	\$104,936,000	\$85,265,000
Group Quarters & Puerto Rico	\$8,975,000	\$13,207,000	\$10,731,000
Subtotal Data Collection	\$125,525,000	\$184,718,000	\$150,090,000
Other Costs	\$29,756,000	\$29,756,000	\$29,756,000
Projected Total Survey Costs	\$155,281,000	\$214,473,000	\$179,846,000
Increase over Mandatory	\$0	\$59,192,000	\$24,565,000

KEY: Current Reliability (Mandatory Methods) reflects 3 million initial annual sample and use of Current Mandatory methods.
 Current Reliability (Voluntary Methods) reflects 3.7 million initial annual sample and use of Standard Voluntary methods.
 Lower Reliability (Voluntary Methods) reflects 3 million initial annual sample and use of Standard Voluntary methods.

5. REFERENCES

- Alexander, C. A. 2002. "Still Rolling: Leslie Kish's 'Rolling Samples' and the American Community Survey." Statistics Canada Survey Methodology, Volume 28, No. 1.
- Asiala, M. 2003. "American Community Survey (ACS) Voluntary Test - Stratification Issues." Internal Census Bureau Memorandum, ACS Voluntary Test Memorandum #8.
- Landreth, A. 2003. "Final Report: Results and Recommendations from Cognitive Interviews with Selected Materials accompanying the American Community Survey." Internal Census Bureau Memorandum, ACS Voluntary Test Memorandum #4.
- Leslie, T. 2003. "Copies of Final Mail Forms for the ACS Voluntary Test." Internal Census Bureau Memorandum, ACS Voluntary Test Memorandum #3.
- McGovern, P. 2003. "Calculation of Data Completeness Rates for the ACS Voluntary Test." Internal Census Bureau Memorandum, ACS Voluntary Test Memorandum #6.

- OMB. 2001. "Statistical Policy Working Paper 31: Measuring and Reporting Sources of Errors in Surveys".
- Raglin, D. 2003. "Calculation of Response Measures for the ACS Voluntary Test." Internal Census Bureau Memorandum, ACS Voluntary Test Memorandum #5.
- Tersine, A. 2003. "Documentation of ACS Voluntary Test Detectable Differences and Additional Results." Internal Census Bureau Memorandum, ACS Voluntary Test Memorandum #7.
- The American Association for Public Opinion Research. 2000. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." Lenexa, Kansas: AAPOR. www.aapor.org.
- Weldon, D. and Miller, D. 2002. "Congress of the United States House of Representatives, Committee on Government Reform Memorandum for the Honorable Frank Wolf, Chairman, Congress, Justice, and State Appropriations Committee".
- U.S. Census Bureau. 2002. "Meeting 21st Century Demographic Data Needs—Implementing the American Community Survey: Demonstrating Survey Quality".

Appendix 1. Summary of Voluntary and Mandatory Differences in Mail Materials

Messages on the Envelope

2003 Voluntary

The American Community Survey
Form Enclosed

**YOUR RESPONSE IS
IMPORTANT TO YOUR
COMMUNITY**

2002 Mandatory

The American Community Survey
Form Enclosed

**YOUR RESPONSE IS
REQUIRED BY LAW**

Messages in the Mailing Package Letters

2003 Voluntary

The U.S. Census Bureau chose your address, not you personally, as part of a randomly selected sample. The Census Bureau is required by U.S. law to keep your answers confidential. Your participation in the survey is important; however, you may decline to answer any or all questions. The enclosed brochure answers frequently asked questions about the survey.

2002 Mandatory

Please take about 40 minutes of your time to assist your community greatly by completing and mailing back your copy of the American Community Survey, as required by law. We are conducting this survey under the authority of Title 13, United States Code, sections 141-193, and 221. That same law protects your privacy—Section 9 requires us to keep all information about you and your household strictly confidential. We may use this information only for statistical purposes. In addition, Title 13 imposes severe criminal sanctions if any U.S. Census Bureau employee violates these provisions. Title 13 also imposes penalties for not responding to the American Community Survey.

Appendix 1. Summary of Voluntary and Mandatory Differences in Mail Materials

p 2 of 2

Message in the Brochure Accompanying the Form

2003 Voluntary

Do I have to answer the questions on the American Community Survey?

Your participation in this voluntary survey is very important to your country and to your community. You may decline to answer any or all questions. The Census Bureau is conducting the survey under the authority of Title 13, United States Code, Section 182. This survey is approved by the Office of Management and Budget. We estimate this survey will take about 38 minutes to complete.

2002 Mandatory

(There was no mention of the mandatory nature of the survey in the brochure).

Appendix 2. Detailed Computation of Cooperation Rates, Weighted Survey Response Rates, Interview Mode Distributions, and Interview Rates

Survey cases can be categorized into four outcomes:

- (1) interviews (including partial interviews with sufficient data);
- (2) eligible cases that are not interviewed (nonrespondents);
- (3) cases of unknown eligibility; and
- (4) cases that are not eligible.

Cases that are determined not to be eligible are not included in the universe for calculation of cooperation or response rates, or interview mode distributions, but they are used in the calculation of interview rates. Ineligible cases include businesses or demolished housing units, and initially selected addresses that are not included in the subsample for personal visit follow-up. Let X = ineligible cases (final, after all attempts). In the ACS, there are no cases with unknown eligibility, so all eligible cases can be coded as one of the following:

Interviews:

- I = Complete interviews (mail, telephone, and personal visit combined)
- I(m) = Complete mail interviews
- I(t) = Complete telephone interviews
- I(p) = Complete personal visit interviews
- P = Partial interviews (mail, telephone, and personal visit combined)
- P(m) = Partial mail interviews
- P(t) = Partial telephone interviews
- P(p) = Partial personal visit interviews

Eligible cases that are not interviewed (nonrespondents):

- R = Refusals (final, after all attempts)
- R(m) = Refusal to the mail attempt
- R(t) = Refusal to telephone attempt
- R(p) = Refusal to personal visit attempt
- NC = Non contacts (final, after all attempts)
- NC(m) = Non contact to mail attempt
- NC(t) = Non contact to telephone attempt
- NC(p) = Non contact to personal visit attempt
- O = Other Noninterviews (final, after all attempts)
- O(m) = Other Noninterviews to mail attempt
- O(t) = Other Noninterviews to telephone attempt
- O(p) = Other Noninterviews to personal visit attempt

These codes are used to produce cooperation and response rates as well as interview mode distributions. Note that for the mail rates, refusals and other noninterviews cannot be distinguished since the reason for nonresponse is unknown. All rates are weighted to reflect the probabilities of selection. The cooperation rates, survey response rates, and interview mode distributions are also weighted to reflect subsampling for personal visit follow-up.

Appendix 2. Detailed Computation of Cooperation Rates, Weighted Survey Response Rates, Interview Mode Distributions, and Interview Rates

p 2 of 4

1. Cooperation Rates

The cooperation rate is defined as the proportion of all cases interviewed of all eligible units ever contacted. We use cooperation rate 2 from the AAPOR standard definitions. (See American Association for Public Opinion Research. 2000, p. 38.) This definition counts partial interviews as respondents. Cooperation rates were calculated for each of the three data collection modes - mail, telephone, and personal visit. The denominator for the mail cooperation rate is all occupied units included in the mailout. The denominators for the telephone and personal visit cooperation rates include all occupied units contacted during follow-up. Specifically,

$$\text{Mail Cooperation Rate} = \frac{I(m) + P(m)}{I(m) + P(m) + R(m) + O(m)} * 100$$

$$\text{Telephone Cooperation Rate} = \frac{I(t) + P(t)}{I(t) + P(t) + R(t) + O(t)} * 100$$

$$\text{Personal Visit Cooperation Rate} = \frac{I(p) + P(p)}{I(p) + P(p) + R(p) + O(p)} * 100$$

Appendix 2. Detailed Computation of Cooperation Rates, Weighted Survey Response Rates, Interview Mode Distributions, and Interview Rates

p 3 of 4

2. Weighted Survey Response Rates

Weighted survey response rates are defined as the number of complete interviews with reporting units divided by the number of eligible reporting units in the sample. We used response rate 6 from the AAPOR standard definitions. (See American Association for Public Opinion Research, 2000, p. 38.) It assumes that there are no cases of unknown eligibility and includes partial interviews as respondents. Specifically,

$$\text{Weighted Survey Response Rate} = \frac{(I + P)}{(I + P) + (R + NC + O)} * 100$$

3. Interview Mode Distributions

Interview mode distributions are defined as the ratio of completed occupied interviews from a specific mode (e.g., mail) to the total completed occupied interviews. Partial interviews are considered respondents. Specifically,

$$\text{Percent Mail} = \frac{I(m) + P(m)}{I + P} * 100$$

$$\text{Percent Telephone} = \frac{I(t) + P(t)}{I + P} * 100$$

$$\text{Percent Personal Visit} = \frac{I(p) + P(p)}{I + P} * 100$$

Appendix 2. Detailed Computation of Cooperation Rates, Weighted Survey Response Rates, Interview Mode Distributions, and Interview Rates

p 4 of 4

4. Interview Rates

Interview rates are defined as the number of complete interviews with reporting units divided by the number of units in the initial (or first phase) sample. The interview rate measures the impact of subsampling and nonresponse on the final number of completed interviews. Specifically,

$$\textit{Interview Rate} = \frac{(I + P)}{(I + P) + (R + NC + O + X)} * 100$$

Appendix 3. Detailed Computation of Data Completeness Rates

For a given item, people or units can be classified into one of three main groups:

- (1) Eligible to respond to the item,
- (2) Not eligible to respond to the item, or
- (3) Unknown eligibility.

All data completeness rates are based on people or units determined to be eligible to respond to an item. A person or unit is not in universe for a particular item without definitive information to determine eligibility to answer the question.

The responses for all eligible people or units can be categorized into one of the following two groups: nonresponse or valid response. Therefore, responses to item x as can be classified as follows:

- N_x = nonresponse to item x due to one of the following reasons: don't know, refused, illegal value, other nonresponse
 V_x = response to item x determined to be valid

The data completeness rate does not measure correctness of the data, only whether or not a valid response was obtained. A valid response doesn't take the consistency of the response or following the proper skip patterns on the questionnaire into account.

The item-level data completeness rate is calculated as the ratio of the number of eligible units or people having a valid response to an item to the total number of units or people eligible to have responded to that item. Therefore, the item-level data completeness rate for item x is computed as follows:

$$\text{Item-Level Data Completeness Rate for Item } x = \frac{V_x}{N_x + V_x} * 100$$

The overall data completeness rate is a measure of the completeness of the entire questionnaire. It is defined as the ratio of the total number of items for which a valid response was given to the total number of items for which a response was required. Specifically,

$$\text{Overall Data Completeness Rate} = \frac{\sum_x V_x}{\sum_x (N_x + V_x)} * 100$$

Appendix 4. Derivation of Revised Sample Size

The full implementation of the ACS calls for a monthly sample of 250,000 addresses or an annual sample of 3 million addresses selected from the approximately 120 million addresses in the country. Thus, the basic sampling weight for each sample housing unit address is roughly 40 ($120,000,000 \div 3,000,000$). Nonrespondents after the telephone follow-up are further subsampled at a rate of 1-in-3, producing a final sampling weight of 120 ($40 * 3$). Incomplete addresses, referred to as unmailables, are further subsampled at a rate of 2-in-3, for a final sampling weight of 60 ($40 * 3/2$).

The variance is a function of the sum of the square of the survey respondents' weights, adjusted up to account for the nonresponding cases. Given that, the increase in variance under a voluntary ACS is due to two factors:

- The reduction in the total number of interviews which reduces the amount of information available to produce sample estimates, and,
- A higher proportion of sample interviews come from the personal visit phase with the largest weights of 120, since more cases are eligible for personal visit follow-up.

Table A gives a distribution of the total number of interviews by data collection mode with their associated weights:

Table A: Total number of interviews by mode

Completed Interviews	Initial sample of 3 million addresses		Sampling Weight
	Mandatory Methods (Percent)	Voluntary Methods (Percent)	
Total Completed Interviews	2,175,000 (100.0)	1,797,000 (100.0)	-----
Mail	1,490,000 (68.5)	942,000 (52.4)	40
Telephone Followup	271,000 (12.5)	337,000 (18.8)	40
Personal Visit Followup – Mailed Cases	344,000 (15.8)	449,000 (25.0)	120
Personal Visit Followup – Nonmailed Cases	70,000 (3.2)	69,000 (3.8)	60

The use of voluntary methods leads to a reduction of about 378,000 interviews or about 13 percent of the initial 3 million sample size. There is also a shift in the distribution of weights. Under voluntary, a much higher proportion of sample interviews carry the largest weight of 120 (25.0 percent versus 15.8 percent). These two factors combined increase the variance and result in estimates with lower reliability.

We estimate that a voluntary ACS with an initial sample size of 3 million would have variances that would be 1.23 times higher (or 23 percent higher) than the variances of a mandatory ACS. Therefore, to maintain the desired level of reliability, under a voluntary ACS, would require an initial sample size of $3,000,000 * 1.23 \approx 3,700,000$ addresses per year. Similarly, a voluntary ACS with an initial sample size of 3 million would have the same reliability as a mandatory ACS with an initial sample of $3,000,000 / 1.23 \approx 2,400,000$ addresses.

Appendix 5. Workload Definitions and Determination

Table 13 (on page 16) shows the detailed annual costs for the ACS by operation under the three designs discussed in this report – (1) initial sample of 3 million using mandatory methods, (2) an initial sample of 3.7 million using voluntary methods to achieve the same reliability, and (3) an initial sample of 3 million using voluntary methods that does not achieve the same reliability.

The workloads for each of the operations included in the cost model are defined in greater detail below. Tables are provided with the parameters used to project the workloads and the expected number of completed interviews. Note that all parameters and universe projections are rounded.

1. Mail

The workload for the mail operation is the universe of cases determined to be eligible for the mailout, or mailable. Sample cases are unmailable if address information is not available on the Census Bureau’s Master Address File or if the address is a post office box (even if additional address information is available). Table A shows the sample, unmailable and mailable parameters and universes, and the expected number of mail interviews for each design. We assume that 4 percent of the sample are unmailable under all three designs and thus 96 percent of the sample each month is the workload for the mail operation. The number of expected mail interviews is about 19 percentage points higher under the mandatory design.

Table A. Mail Parameters and Workloads - Annual Projections

	Current Reliability (Mandatory Methods)		Current Reliability (Voluntary Methods)		Lower Reliability (Voluntary Methods)	
	Parameter	Universe	Parameter	Universe	Parameter	Universe
Sample		3,000,000		3,700,000		3,000,000
Unmailables	4.0% of sample	120,000	4.0% of sample	148,000	4.0% of sample	120,000
Mailables - Workload	96.0% of sample	2,880,000	96.0% of sample	3,544,000	96.0% of sample	2,880,000
Mail interviews	51.7% of workload	1,489,000	32.7% of workload	1,159,000	32.7% of workload	942,000

2. Telephone follow-up

The workload for telephone follow-up is the universe of cases eligible for mailout for which a mail return was not received after the first month of data collection and for which a telephone number was available. Table B shows the telephone follow-up workloads and expected interviews for these three designs. We assume that under mandatory methods about 35 percent of the sample is sent to telephone follow-up. Under voluntary, that rate increases to over 40 percent due to the additional nonresponse to the mail. Despite the drop in cooperation, a greater percent of the telephone follow-up workload resulted in a completed interview. Under

Appendix 5. Workload Definitions and Determination

p 2 of 3

voluntary methods, a higher percentage of the telephone numbers reached eligible households. The rate of completed interviews is about 2 percentage points higher for the voluntary designs.

Table B. Telephone Follow-up Parameters and Workloads - Annual Projections

	Current Reliability (Mandatory Methods)		Current Reliability (Voluntary Methods)		Lower Reliability (Voluntary Methods)	
	Parameter	Universe	Parameter	Universe	Parameter	Universe
Sample		3,000,000		3,700,000		3,000,000
Workload	35.0% of sample	1,050,000	40.2% of sample	1,485,000	40.2% of sample	1,206,000
Telephone interviews	25.8% of workload	271,000	27.9% of workload	415,000	27.9% of workload	337,000

3. Personal visit follow-up

The workload for personal visit follow-up comes from two sources:

- A sample of 1-in-3 is selected from the cases determined to be mailable and for which a completed mail or telephone interview was not received.
- A sample of 2-in-3 is selected from the cases determined to be unmailable

Table C shows the annual personal visit follow-up workload for each design based on the combination of mailable and unmailable cases. The workload is approximately 67 percent of the unmailable universe plus about 33 percent of the mailable universe without a mail or telephone follow-up response. We assume that under the Current Mandatory design about 16 percent of the sample is sent to personal visit follow-up. Due to decreases in mail and telephone response, this rate is about 21 percent under both voluntary designs. The percent of personal visit follow-up interviews expected is about 4 percentage points higher under the Current Mandatory design.

Table C. Personal Visit Follow-up Parameters and Workloads - Annual Projections

	Current Reliability (Mandatory Methods)		Current Reliability (Voluntary Methods)		Lower Reliability (Voluntary Methods)	
	Parameter	Universe	Parameter	Universe	Parameter	Universe
Sample		3,000,000		3,700,000		3,000,000
Total Workload	16.0% of sample	480,000	21.1% of sample	777,000	21.1% of sample	632,000
Total Personal Visit Interviews	86.3% of workload	414,000	82.1% of workload	638,000	82.1% of workload	518,000

4. Group quarters

Starting in FY 2005, 4,000 sample units (for example, rooms, beds, or people) in living quarters other than housing units will be included in the ACS sample each month (48,000 annually).

Appendix 5. Workload Definitions and Determination

p 3 of 3

5. Puerto Rico

Starting in FY 2005, 3,000 sample cases will be included in the ACS sample each month (36,000 annually).

Appendix 6. Cost Model

The cost model is broken down into two major components - costs for data collection and all other costs. Data collection costs include the costs for mail, telephone follow-up, personal visit follow-up, group quarters interviewing, and Puerto Rico. The components of costs for each of these operations is defined below:

Mail: We have projected the per unit mailing cost of a mailable case at \$12.50. This includes printing and mailout costs associated with all mailings (prenotice letter, initial and second mailing packages, reminder card), check-in and data capture of mail returns, telephone questionnaire assistance support, edit review of mail returns, and telephone follow-up of mail returns with incomplete or inconsistent data. Under voluntary, the second mailing operation will cost more than under the mandatory design but the data entry and telephone follow-up of mail returns with incomplete or inconsistent data will cost less than under the mandatory design, offsetting the higher costs for the second mailing operation.

Telephone follow-up: We have projected the per unit cost of a telephone follow-up case at \$15. This includes the production costs of the call center staff. Costs are incurred for cases with bad telephone numbers and refusals, and for completed interviews. This estimate includes certain infrastructure costs of training, equipment, and staff and workload management.

Personal visit follow-up: We have projected the cost of personal visit follow-up at \$135 per case. Included are travel and interviewing costs. This estimate includes certain infrastructure costs of equipment, training, staff salary, and workload management.

For this report, we assumed the per unit costs for a voluntary and mandatory survey were equivalent. We use the following basic formula to calculate the variable data collection costs:

Cost = \$12.50M + \$15T + \$135P for

M = mail workload

T = telephone follow-up workload

P = personal visit follow-up workload

We used the workload estimates described in Appendix 5 in combination with these per unit costs to estimate the data collection costs for the ACS under three assumptions - mandatory, voluntary (3 million sample) and voluntary (3.7 million sample).

Group quarters and Puerto Rico: We have projected the cost of data collection for group quarters interviewing and Puerto Rico at \$8,975,000 under a mandatory design. Cost increases under voluntary were projected to be proportional to the increases seen in the mail, telephone and personal visit follow-up operations stateside. These projections are our best estimate but may overstate or understate the true increase.

Other Costs: We project that the cost for all other support activities for the ACS is \$29,756,000 across designs. These costs include headquarters staff costs to design, conduct, monitor,

Appendix 6. Cost Model

p 2 of 2

research, and evaluate the ACS operations. Included also are costs to process, tabulate, review, and disseminate ACS data annually. Finally, the costs include staff costs to support, develop and maintain the automated data collection instruments used for telephone and personal visit follow-up, and costs to control work sent to telephone and personal visit follow-up interviewers monthly.