

SUPPORTING STATEMENT
U.S. Department of Commerce
U.S. Census Bureau
ACS Messaging Benchmark and Refinement Study
OMB Control No. 0607-0760

This Supporting Statement provides additional information regarding the Census Bureau's request for processing of the proposed information collection. The numbered questions correspond to the order shown on the Office of Management and Budget Form 83-I, "Instructions for Completing OMB Form 83-I."

B. Collections of Information Employing Statistical Methods

1. Universe and Respondent Selection

The universe for this study is US adults (18 years or older) who generally handle the mail for their household, in order to understand the attitudes of those that are most likely to interact with the ACS mail package. According to ACS estimates there were approximately 131 million households in the United States in 2012 (Olson, 2013). This quantitative telephone study uses a stratified sample design of landline and cellphones to evaluate and refine the most effective messages for use in the ACS mail package and other ACSO communications efforts. Within households, we will screen for an adult who generally handles the mail and ask a series of questions including six of eleven messages. The study will consist of two phases: a Benchmark phase and a Refinement phase. Both phases will consist of $n=1,000$ completed interviews among adult (18+) United States residents who generally handle the mail for their household.

In the Benchmark survey, we will use these interviews to compare eleven (11) messages to determine which messages respondents find believable and would make them more likely to complete the American Community Survey. Respondents will hear a random selection of six messages with subsequent questions measuring their responses to those messages. As a result, we expect to measure a Cohen's d 0.28 difference between messages on a five-point scale between the means of messages with 80% power and family-wise $\alpha=.05$ using a Bonferroni correction for multiple comparisons. Effect sizes of this magnitude are generally considered to be between small ($d\sim 0.2$) and medium ($d\sim 0.5$) in size, which is appropriate for our analysis purposes (Cohen 1992).

The same sampling methodology will be used for both the Benchmark and Refinement phase. We will de-duplicate the randomly selected landline and cell phone numbers for the Refinement Phase against the sample used in the Benchmark Phase to ensure that

no respondents are called for both phases.

2. Procedures for Collecting Information

Among the landline interviews, the sample frame will be developed in two stages. First, counties will be stratified based on their 2012 ACS self-response rate. Then, telephone numbers will be randomly generated using known exchanges that are in those strata. As typical with geographic-based RDD frames, only exchanges with at least one listed household per hundred numbers will be included in the sample frame.

Stratification

The research team will use a geographic stratification method to ensure that low-response, medium-response, and high-response areas are properly represented in the sample.

Geographic stratification will be based on ACS self-response rates. The ACSO has operational and population data for each of the 3,142 counties in the United States (and similar geographic entities such as parishes and independent cities). These counties will be ranked in order of their 2012 ACS self-response rates, and divided into three strata with equally sized populations (approximately 102 million people each).

We will conduct n=250 landline interviews respectively within the top-third, middle-third, and bottom-third of U.S. counties based on the ACS self-response rates, as illustrated in the table below:

Frame	Description	Population within strata	# of Counties within strata	Range of ACS self-response rates (2012)
Landline (n=750 /phase)	Top-third of US population, based on ACS response rate, by county	~102,124,000	1,570 counties	100% - 64.7%
	Middle-third of US population, based on ACS response rate, by county	~102,231,000	913 counties	64.7% - 55.9%
	Bottom-third of US population, based on ACS response rate, by county	~102,248,000	660 counties	55.9% - 0.0%
Cell Phone (n=250 /phase)	Randomized based on known cell phone exchanges nationally (no self-response rate targeting)	278,000,000 (estimated*)	National	N/A

*According to the Pew Research Center, 91% of US adults have cell phones (Rainie, 2013).

Sample Selection

Supporting Statement Part B - ACS Messaging Benchmark and Refinement Study
 OMB Control Number 0607-0760

The second stage of developing the sample frame will be generating the list of telephone numbers to contact. Numbers will be selected using Random Digit Dialing (RDD) telephone sampling, which has been used for decades to create a representative sample of the US population. RDD offers excellent coverage of any designated area without the potential biases of other methodologies. As opposed to list-based sampling, which by definition does not include every household in a desired area, RDD generates every possible number in an exchange – including new movers and unlisted numbers.

To produce numbers for the RDD landline sample, the first six digits dialed (area code + exchange) will be determined based on the high, medium, and low stratifications of ACS response rates. The final four digits will be generated randomly. For telephone exchanges that are in multiple strata, we will make an effort to determine which stratum has a greater number of numbers on that exchange, in which case the exchange will be assigned to that stratum. If there cannot be a determination, the exchange will be assigned at random to a single stratum.

Because we will be using RDD, a significant portion of the numbers may not be operable– they may be commercial, fax, or disconnected numbers. When possible, non-household numbers will be identified and removed before contacting.

We assume that the low ACS self-response counties will have a lower telephone survey response rate, and the high-response strata will have a higher telephone response rate. Our initial sample frames for each stratum will reflect this assumption. Therefore, the high-response strata frame will begin with approximately n=15,250 landline numbers, the middle-response strata will begin with approximately n=17,500, and the low-response strata will begin with n=19,750. These sample sizes incorporate estimates about the number of RDD telephone numbers that are answered by households, agree to participate, qualify for and complete the interview within a tolerance to ensure adequate sample. The following table provides a summary of the telephone number sampling:

Frame	Strata description	Sample Phone Numbers
Landline (n=750 /phase)	Top-third of US population, based on ACS response rate, by county	15,250 numbers
	Middle-third of US population, based on ACS response rate, by county	17,500 numbers
	Bottom-third of US population, based on ACS response rate, by county	19,750 numbers
Cell Phone	Randomized based on known cell	20,000 numbers

Supporting Statement Part B - ACS Messaging Benchmark and Refinement Study
OMB Control Number 0607-0760

(n=250 /phase)	phone exchanges nationally (no self-response rate targeting)	
----------------	--	--

While the distribution in each stratum will vary slightly with the total population, we anticipate starting with a total of approximately n=52,500 RDD landline numbers (after removing disconnected, fax, and commercial numbers using computer database software). We assume that the low ACS self-response counties will have a lower telephone survey response rate, and the high response strata will have a higher telephone response rate. Our initial sample frames for each stratum will reflect this assumption.

Using a similar RDD methodology, cell phone interviews will be targeted to a sampling of all known national cell exchanges nationally. The cell phone RDD sample will be randomized to guard against potential regional or demographic bias. This separate sampling frame for cell phone interviews helps us find households that do not have a landline and are “cell phone only.” We anticipate starting with a total of approximately 20,000 RDD cellphone numbers (sample ratio 80:1).

Phone Interviews

Data will be collected through quantitative interviews. The interview begins with an introduction and screening questions to ensure that each respondent is an adult in the household who generally handles the mail. Then the interview turns to questions about the awareness of the ACS; attitudes towards the federal government, including the Census; concerns about the intrusiveness of the ACS; and messages regarding participation in the ACS. The interview concludes with demographic questions.

Up to eight (8) attempts will be made to contact each household in the sample frame. Using area code information, calls will be directed to households during the weeknight evening hours in their particular time zone.

The survey may be completed in English or Spanish. All initial calls will be made in English. Should interviewers identify Spanish-speaking households who indicate a preference in conducting the interview in Spanish, they will call back with a Spanish-language interviewer and questionnaire.

The survey will be programmed using computer aided telephone interviewing (CATI) software, including skip patterns and constrained responses. All interviewers are trained in quantitative interviewing and have experience with telephone interviews with the general public.

Weighting

While the survey will not be used to make estimates of the target population as a whole, the sample will be weighted to ensure that the findings are not unduly influenced by sample imbalances in demographic characteristics such as race, age, and gender.

The target audience for the survey is U.S. adults that generally handle the mail. As reliable demographic estimates of the population who handles the mail are not available, we will use householder (head-of-household) data from the Current Population Survey (CPS) as a proxy for the sample weights. This is not a perfect proxy for the weights, yet it provides a reasonable framework to represent the adults who generally handle the mail for their household.

There is one notable adjustment between CPS householder data and the survey weighting. Gender weights in the Benchmark survey are constructed by combining the number of householders living in non-family households or in family households where no spouse is present for each gender. In addition, married couples living in the same household are considered equally likely to have male or female who handles the mail, so as not to overweight the proportion of married families that identify the male as the householder for the family. This is summarized in the following table:

Number of Householders by gender and family status (in thousands)			
Householder Family Status	Total	Male	Female
Householder not in family household <i>(includes living alone or with nonrelatives)</i>	41,558	19,747	21,810
Householder in family without spouse <i>(includes married spouse absent, widowed, divorced, separated, or never married)</i>	21,699	6,230	15,469
Married with spouse present <i>(* for weighting purposes, married spouses in the same household are considered equally likely to generally handle the mail)</i>	59,204	29,602*	29,602*
Total Householders	122,460	55,579	66,881
Percentage	100%	45%	55%

(Source: CPS 2013 Annual Social and Economic Supplement, Tables A1 and A2)

The research team will use a Random Iterative Method to conduct weighting. Cases with unknown values for particular values, because the respondent answered “don’t know” or refused to answer the question, will be assumed to be unweighted (i.e., weight of 1.0) for that particular item and iteration. The following table details the target demographic weights for the survey:

Demographic Targets for Weighting

Supporting Statement Part B - ACS Messaging Benchmark and Refinement Study
 OMB Control Number 0607-0760

Category	Characteristic	Target Percentage
Gender	Male	45%
	Female	55%
Race	White, non-Hispanic	72%
	White, Hispanic	11%
	Black	12%
	Asian	4%
	All Others	1%
Age	18-29	13%
	30-44	26%
	45-64	38%
	65+	23%
<i>(Source: CPS 2013 Annual Social and Economic Supplement, Tables A1 and A2)</i>		

Data Analysis

Our data analysis approach is designed to identify which messages are most effective at increasing likelihood to participate in the ACS survey.

In the Benchmark phase, the goal is to help determine which messages should be explored further with the subsequent Refinement phase. This further exploration may include A/B testing of message variations, removing less effective messages, or making changes to the initial messages. The research team will decide what messages to explore further based on findings from the Benchmark study as well as other on-going qualitative research projects, such as the deliberative focus groups, key informant interviews, and mental modeling interviews with Census Bureau staff.

The two message variables we are most interested in are the believability of the message (Metric A; “very believable to “very unbelievable”) and likelihood to respond (Metric B; “much more likely to respond” to “much less likely to respond”). Our data analysis approach will compare messages using these two variables in both statistical and non-statistical ways. In the Benchmark phase, each respondent will hear six of the eleven messages. The order will be randomized to address potential order bias and learning effects.

First, we will visually compare messages to identify “big picture” differences between the eleven messages. In particular, we will look at the messages by the superlative, “top box” answer choices (i.e. “very believable” and “much more likely to participate”) to observe which messages have strong impacts among the largest number of people. We will also look at the results by various demographic and attitudinal crosstabs in a non-statistical way. Potentially significant findings can be statistically tested, though the survey instrument is not designed to provide statistical precision for various

demographic sub-groups.

Second, we will test whether the differences between messages are statistically significant using Tukey's homogeneous subsets testing. We will convert the categorical answer choices (i.e., "much more likely to participate" to "much less likely to participate") into a five-point quantitative scale. Using a sample of 545 randomized observations for each message, we expect to measure a Cohen's d 0.28 difference between the average score for messages on a 5-point scale with 80% power and family-wise $\alpha=.05$, using Bonferroni correction for multiple comparisons.

Third, we will analyze changes in the attitude questions before and after the message section to look for statistical relationships between exposure to particular messages and attitudes towards participation in the ACS. We will use a logistic regression method using the difference from pre- to post- likelihood to participate measures as the dependent variable for the model. We will run the regression model using whether participants heard each of the eleven messages as dummy variables (with 1 being they heard the message and a 0 being they did not hear the message). For messages that were not heard or where respondents did not answer the question, a value of 0 will be used in the regression. We will explore whether the fit of the model can be improved with demographic control variables that previous research has found to predict participation in Census Bureau surveys such as age, race, gender, household income, or education (see Bates & Mulry, 2009).

Finally, we will analyze the drilldown on intrusiveness and privacy section that is heard by respondents who are highly distrustful of the government. This analysis will focus on the differences between drilldown statements to identify which approaches to address intrusion and privacy are most likely to increase trust among this naturally distrustful segment of the population.

Because the drilldown analysis is being conducted with a sub-set of the total number of interviews, it will have less statistical precision than the overall analysis. Based on the proportion of respondents likely to be part of the "cynical" or "suspicious" CBAMS II mindsets, we anticipate that approximately 25% of respondents (or $n=250$) will qualify to hear the distrustful drilldown questions. Using a T-test to compare the means between scores, we expect to measure a Cohen's d 0.35 difference between the average score for messages on a 5-point trust scale with 80% power and $\alpha=.10$. Effect sizes of this type are generally considered to be between small and medium effects (Cohen 1992).

Based on this analysis, the research team will:

Supporting Statement Part B - ACS Messaging Benchmark and Refinement Study
OMB Control Number 0607-0760

- Report on key communications findings, including guidance on effective topics, themes, and tones for increasing interest in the ACS process and support of its mission
- Identify messages that effectively address concerns about privacy, intrusiveness, and harassment among survey participants who had distrustful or skeptical attitudes towards government
- Produce communications recommendations that can broadly inform other communications, including talking points, press releases, and the website
- Recommend words, phrases, topics, and tone to use for further research in the Mail Package Assessment.

As this study will be conducted under the CLMSO's Generic Clearance for Data User and Customer Evaluation, this study will not be used to draw inferences regarding the country's population at-large and will not be used to publish any official statistical estimates. As part of the study's iterative design, the questionnaire for the Refinement Study will be revised based on findings from the Benchmark Study and other ongoing Census Bureau research to identify the most effective messages about ACS participation, among a variety of stakeholders.

3. Methods to Maximize Response

As this purpose of this research is to help inform the Census Bureau's internal decision-making about messages to respondents, the study does not need the statistical rigor that OMB typically requires for findings that will produce statistical estimates of the population or public dissemination. We anticipate a response rate between 2% (APPOR 1) and 12% (APPOR 3). This study incorporates several design features to improve response. An RDD-sampling approach provides minimum coverage omissions by providing the broadest possible response universe on the telephone. Interviewers will attempt up to eight (8) contacts, in English or Spanish, to minimize non-response errors. The survey instrument has been cognitive tested to improve cooperation rates and to ensure clarity, brevity, relevance, and user-friendliness.

4. Test of Procedures or Methods

Census Bureau statisticians who routinely design census and survey other research studies review all research methodology and documentation. Staff will systematically monitor data collection procedures in order to identify ways to reduce burden, streamline processing, and assure quality data.

The questionnaire was also cognitively tested with nine respondents from a variety of

Supporting Statement Part B - ACS Messaging Benchmark and Refinement Study
OMB Control Number 0607-0760

backgrounds. Following testing, the contractor developed seven recommended revisions to improve the data collection instrument. The report from the Cognitive Testing is attached (see Attachment D, Cognitive Interviewing Report).

5. Contacts for Statistical Aspects and Data Collection

Consultants outside of the Census Bureau are listed below.

Kiera McCaffrey

Reingold
202.559.4436
kmccaffrey@reingold.com

Sam Hagedorn

Penn Schoen Berland
202.962.3056
shagedorn@ps-b.com

Jack Benson

Reingold
202.333.0400
jbenson@reingold.com

Robert Green

Penn Schoen Berland
202.962.3049
rgreen@ps-b.com

The data will be collected by contractor Penn Schoen Berland.

Attachments

- A – References and Works Cited
- B – Cognitive Interview Report
- C -- Benchmark Survey Questionnaire (English)
- D – Benchmark Survey Questionnaire (Spanish)